Jan-Feb.
1960
D. Radio
Propagation
64D
No. 1

SH11.2

Volume 64D

Number 1

Jan.–Feb. 1960

RECEIVED
JAN 18 1960
MIAMI UNIVERSITY
LIBRARY

Journal of Research

of the

NATIONAL BUREAU OF STANDARDS

D. RADIO PROPAGATION

Journal of Research

of the

National Bureau of Standards

D. Radio Propagation

JAN.–FEB. • 1960

VOLUME • 64D

NUMBER • 1

Editor: James R. Wait

Associate Editors: Thomas N. Gautier, Jack W. Herbstreit, C. Gordon Little, Alvin G. McNish, Richard C. Kirby

IRE Advisors: D. G. Fink, K. M. Siegel

Publication dates:
Jan. 15; Mar. 15; May 15; July 15; Sept. 15; Nov. 15, 1960

JOURNAL OF RESEARCH

The National Bureau of Standards Journal of Research reports research and development in the fields of activity shown at right. Also included from time to time are survey articles and compilations of information on topics closely related to the Bureau's scientific and technical program.

The Journal is published in four separate sections as follows:

A. Physics and Chemistry

Contains papers of interest primarily to scientists working in these fields. Issued six times a year. Annual subscription: domestic, $4.00; foreign, $4.75.

B. Mathematics and Mathematical Physics

Presents studies and compilations designed mainly for the mathematician and the theoretical physicist. Issued quarterly. Annual subscription: domestic, $2.25; foreign, $2.75.

C. Engineering and Instrumentation

Reports research and development results of interest chiefly to the engineer and the applied scientist. Issued quarterly. Annual subscription: domestic, $2.25; foreign, $2.75.

D. Radio Propagation

Reports work in radio propagation, upper atmospheric physics, and communications. Issued six times a year. Annual subscription: domestic, $4.00; foreign, $4.75.

Functions and Activities

The functions of the National Bureau of Standards are set forth in the Act of Congress, March 3, 1901, as amended. These include the development and maintenance of the national standards of measurement and the provision of means and methods for making measurements consistent with these standards; the determination of physical constants and properties of materials; the development of methods and instruments for testing materials, devices, and structures; advisory services to government agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; and the development of standard practices, codes, and specifications. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services, and various consultation and information services. The Bureau also serves as the Federal technical research center in a number of specialized fields. The scope of activities of the National Bureau of Standards is suggested in the following listing of the divisions and sections engaged in technical work.

Washington, D.C.

Electricity and Electronics. Resistance and Reactance. Electron Devices. Electrical Instruments. Magnetic Measurements. Dielectrics. Engineering Electronics. Electronic Instrumentation. Electrochemistry.

Optics and Metrology. Photometry and Colorimetry. Optical Instruments. Photographic Technology. Length. Engineering Metrology.

Heat. Temperature Physics. Thermodynamics. Cryogenic Physics. Rheology. Molecular Kinetics. Free Radicals Research.

Atomic and Radiation Physics. Spectroscopy. Radiometry. Mass Spectrometry. Solid State Physics. Electron Physics. Atomic Physics. Neutron Physics. Radiation Theory. Radioactivity. X-ray. High Energy Radiation. Nucleonic Instrumentation. Radiological Equipment.

Chemistry. Organic Coatings. Surface Chemistry. Organic Chemistry. Analytical Chemistry. Inorganic Chemistry. Electrodeposition. Molecular Structure and Properties of Gases. Physical Chemistry. Thermochemistry. Spectrochemistry. Pure Substances.

Mechanics. Sound. Mechanical Instruments. Fluid Mechanics. Engineering Mechanics. Mass and Scale. Capacity, Density, and Fluid Meters. Combustion Controls.

Organic and Fibrous Materials. Rubber. Textiles. Paper. Leather. Testing and Specifications. Polymer Structure. Plastics. Dental Research.

Metallurgy. Thermal Metallurgy. Chemical Metallurgy. Mechanical Metallurgy. Corrosion. Metal Physics.

Mineral Products. Engineering Ceramics. Glass. Refractories. Enameled Metals. Constitution and Microstructure.

Building Technology. Structural Engineering. Fire Protection. Air Conditioning, Heating, and Refrigeration. Floor, Roof, and Wall Coverings. Codes and Safety Standards. Heat Transfer. Concreting Materials.

Applied Mathematics. Numerical Analysis. Computation. Statistical Engineering. Mathematical Physics.

Data Processing Systems. SEAC Engineering Group. Components and Techniques. Digital Circuitry. Digital Systems. Analog Systems. Applications Engineering.

• Office of Basic Instrumentation. • Office of Weights and Measures.

Boulder, Colorado

Cryogenic Engineering. Cryogenic Equipment. Cryogenic Processes. Properties of Materials. Gas Liquefaction.

Radio Propagation Physics. Upper Atmosphere Research. Ionosphere Research. Regular Prediction Services. Sun-Earth Relationships. VHF Research. Radio Warning Services. Airglow and Aurora. Radio Astronomy and Arctic Propagation.

Radio Propagation Engineering. Data Reduction Instrumentation. Radio Noise. Tropospheric Measurements. Tropospheric Analysis. Propagation-Terrain Effects. Radio-Meteorology. Lower Atmosphere Physics.

Radio Standards. High-Frequency Electrical Standards. Radio Broadcast Service. Radio and Microwave Materials. Electronic Calibration Center. Microwave Circuit Standards.

Radio Communication and Systems. Low Frequency and Very Low Frequency Research. High Frequency and Very High Frequency Research. Modulation Systems. Antenna Research. Navigation Systems. Systems Analysis. Field Operations.

U.S. DEPARTMENT OF COMMERCE
Frederick H. Mueller, Secretary

NATIONAL BUREAU OF STANDARDS
A. V. Astin, Director

Order all publications from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C.

Use of funds for printing this publication approved by the Director of the Bureau of the Budget (June 24, 1958).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Effect of Antenna Size on Gain, Bandwidth, and Efficiency[1]

Roger F. Harrington[2]

(June 29, 1959)

A theoretical analysis is made of the effect of antenna size on parameters such as gain, bandwidth, and efficiency. Both near-zone and far-zone directive gains are considered. It is found that the maximum gain obtainable from a broad-band antenna is approximately equal to that of the uniformly illuminated aperture. If higher gain is desired, the antenna must necessarily be a narrow-band device. In fact, the input impedance becomes frequency sensitive so rapidly that, for large antennas, no significant increase in gain over that of the uniformly illuminated aperture is possible. Also, if the antenna is lossy, the efficiency falls rapidly as the gain is increased over that of the uniformly illuminated aperture.

1. Introduction

As a practical matter, the maximum directive gain (directivity) of an antenna depends upon its physical size compared to wavelength. The uniformly illuminated aperture type of antenna has been found to give a higher gain in practice than other antennas, at least for large apertures. However, the uniformly illuminated aperture does not represent a theoretical limit to the gain. Higher directive gains appear to be possible, but analyses of projected "supergain" antennas reveal extreme frequency sensitivity at best, excessive losses at worst. This paper gives a theoretical treatment to the general problem, from which quantitative bounds to antenna performance may be obtained. The analysis considers both the near-zone and the far-zone gain of antennas.

Let the spherical coordinate system be defined as in figure 1. The directive gain as a function of distance from an antenna is defined as the ratio of the maximum density of outward-directed power flux to the average density. In equation form this is

$$G(\beta r)=\frac{4\pi r^2 \mathrm{Re}(S_r)_{\max}}{\mathrm{Re}(P)}, \tag{1}$$

where S_r is the radial component of the complex Poynting vector at a distance r, and P is the outward-directed complex power over a sphere of radius r.

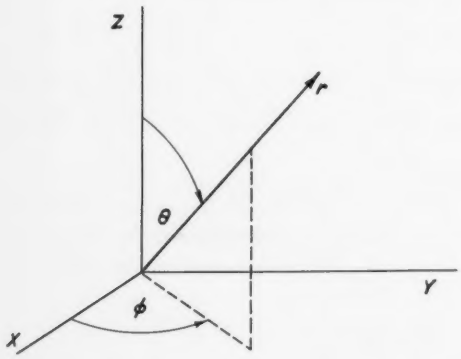


FIGURE 1. *The spherical coordinate system.*

[1] This work has been supported by General Electric Company, subcontract No. 500-7119
[2] Department of Electrical Engineering, Syracuse University, Syracuse, N.Y.

The field external to a sphere containing all sources can be expanded in terms of spherical wave functions. The general form is [1] [3]

$$\boldsymbol{E}=-\nabla\times(\mathbf{r}\psi)+\frac{1}{j\omega\epsilon}\nabla\times\nabla\times(\mathbf{r}\hat{\psi})$$
$$\boldsymbol{H}=\nabla\times(\mathbf{r}\hat{\psi})+\frac{1}{j\omega\mu}\nabla\times\nabla\times(\mathbf{r}\psi) \tag{2}$$

where $\boldsymbol{r}$ is the radius vector from the origin and

$$\psi=\sum_{m,n} A_{mn}h_n^{(2)}(\beta r)P_n^m(\cos\ \theta)\ \cos\ (m\phi+\alpha_{mn})$$
$$\hat{\psi}=\sum_{m,n} B_{mn}h_n^{(2)}(\beta r)P_n^m(\cos\ \theta)\ \cos\ (m\phi+\hat{\alpha}_{mn}) \tag{3}$$

where A_{mn}, B_{mn}, α_{mn}, and $\hat{\alpha}_{mn}$ are coefficients which do not depend on r, θ, or ϕ, and $\beta=2\pi/$wavelength. From the above formulas for the field can be calculated the power over a sphere of radius r

$$\mathrm{Re}(P)=\frac{4\pi}{\beta^2}\sum_{m,n}\frac{n(n+1)(n+m)!}{\epsilon_m(2n+1)(n-m)!}\left(\frac{1}{\eta}|A_{mn}|^2+\eta|B_{mn}|^2\right) \tag{4}$$

where ϵ_m is Neumann's number and $\eta\cong 120\pi$. The radial component of the power flux density in the $\theta=0$ direction is

$$S_r=E_xH_y^*-E_yH_x^* \tag{5}$$

where

$$E_x=\sum_n\frac{n(n+1)}{2\beta r}\left[A_{1n}\sin\alpha_{1n}F_n(\beta r)-j\eta B_{1n}\cos\hat{\alpha}_{1n}F_n'(\beta r)\right]$$
$$E_y=\sum_n\frac{n(n+1)}{2\beta r}\left[A_{1n}\cos\alpha_{1n}F_n(\beta r)+j\eta B_{1n}\sin\hat{\alpha}_{1n}F_n'(\beta r)\right] \tag{6}$$
$$H_x=\sum_n\frac{n(n+1)}{2\beta r}\left[-B_{1n}\sin\hat{\alpha}_{1n}F_n(\beta r)+\frac{j}{\eta}A_{1n}\cos\alpha_{1n}F_n'(\beta r)\right]$$
$$H_y=\sum_n\frac{n(n+1)}{2\beta r}\left[-B_{1n}\cos\hat{\alpha}_{1n}F_n(\beta r)-\frac{j}{\eta}A_{1n}\sin\alpha_{1n}F_n'(\beta r)\right]$$

and

$$F_n(\beta r)=\beta r\,h_n^{(2)}(\beta r). \tag{7}$$

The coordinate axes are to be oriented so the maximum radiation is in the $\theta=0$ direction. Then the S_r of (5) is the $(S_r)_{\max}$ of (1).

Note that α_{mn} and $\hat{\alpha}_{mn}$ do not enter into the formula for power, (4). Also, from the symmetry of (5) and (6) it is evident that the density of power flux in the $\theta=0$ direction is independent of α_{1n} and $\hat{\alpha}_{1n}$. Hence, they may be chosen in any convenient manner. In particular, let $\alpha_{1n}=\pi/2$, $\hat{\alpha}_{1n}=\pi$, which give a field linearly polarized in the x direction. With this choice, the gain (1) becomes

$$G=\frac{\mathrm{Re}[(\sum_n a_nF_n+jb_nF_n')(\sum_n b_nF_n+ja_nF_n')^*]}{4\sum_{m,n}\frac{n(n+1)(n+m)!}{\epsilon_m(2n+1)(n-m)!}(|A_{mn}|^2+|\eta B_{mn}|^2)} \tag{8}$$

[3] Figures in brackets indicate the literature references at the end of this paper.

where

$$a_n = n(n+1)A_{1n} \qquad b_n = \eta n(n+1)B_{1n}. \tag{9}$$

Equation (8) is a general formula for directive gain.

2. Maximum Gain

If all orders of spherical wave functions are permitted there is no limit to the gain of an antenna. However, a definite limit to the gain exists if wave functions are restricted to orders $n \leq N$. Only the A_{1n} and B_{1n} contribute to the numerator of (8), so the gain can be increased by setting

$$A_{mn} = B_{mn} = 0 \qquad m \neq 1. \tag{10}$$

Furthermore, the gain formula is symmetrical in a_n and b_n so the maximum gain will exist under the condition

$$a_n = b_n. \tag{11}$$

Equation (8) has now been reduced to

$$G = \frac{\mathrm{Re}\left[\left(\sum_{n=1}^{N} a_n u_n\right)\left(\sum_{n=1}^{N} a_n u_n\right)^*\right]}{4\sum_{n=1}^{N} |a_n|^2 \frac{1}{2n+1}} \tag{12}$$

where

$$u_n(\beta r) = F_n(\beta r) + j F_n'(\beta r). \tag{13}$$

The numerator of (12) can now be increased without changing the denominator by setting

$$\text{phase } (a_n) = -\text{phase } (u_n) \tag{14}$$

in which case (12) becomes

$$G = \frac{\left(\sum_{n=}^{N} |a_n|\,|u_n|\right)^2}{4\sum_{n=1}^{N} |a_n|^2 \frac{1}{2n+1}}. \tag{15}$$

Finally, the $|a_n|$ are adjusted for maximum gain by requiring $\partial G/\partial |a_i| = 0$ for all a_i. The result is

$$G(\beta r)_{\max} = \tfrac{1}{4}\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2 \tag{16}$$

which is the maximum directive gain obtainable using wave functions of order $n \leq N$. The maximization procedure also results in the relationship

$$\left|\frac{a_n}{a_i}\right| = \frac{(2n+1)|u_n|}{(2i+1)|u_i|}. \tag{17}$$

From (13) it follows that

$$|u_n(\beta r)|^2 = |F_n'(\beta r)|^2 + |F_n(\beta r)|^2 + 2. \tag{18}$$

As $\beta r \to \infty$, $|u_n|^2 \to 4$, so in the far zone (16) reduces to

$$G(\infty)_{\max} = \sum_{n=1}^{N}(2n+1) = N^2 + 2N \tag{19}$$

which has been previously published [1].

The $|u_n|^2$ functions, which enter into (16) and also into later formulas, are shown in figure 2. The maximum directive gain for various βr and various N has been calculated. Figure 3 shows the ratio of the near-zone gain to the far-zone gain for several N. Note that the maximum near-zone gain is essentially the same as the maximum far-zone gain unless $\beta r < N$. Keep in mind that the excitation of the antenna is changed as βr is varied so that it is always adjusted for maximum gain at the given radius. One can think of the antenna as being focused at the distance r.

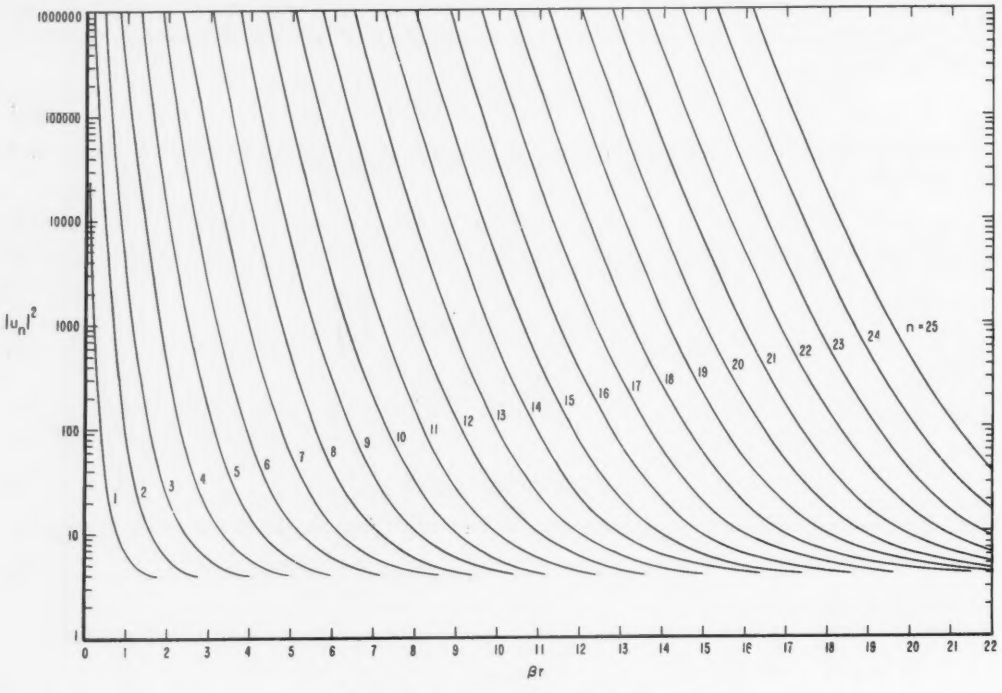


FIGURE 2. *The functions* $|u_n(\beta r)|^2$.

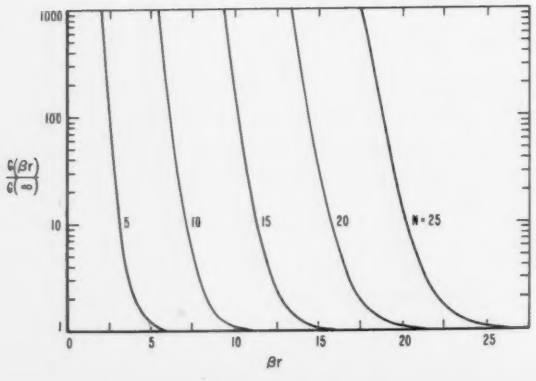


FIGURE 3. *Ratio of the maximum gain at a distance* r *to the maximum gain at infinity, using wave functions of order* $n \leq N$.

3. Quality Factor

So far the antenna structure has not been mentioned. An ideal loss-free antenna of radius R is defined as one having no energy storage $r<R$, which is the same as Chu's definition [2]. The quality factor

$$Q=\begin{cases}\dfrac{2\omega W^{\text{elect}}}{\text{Re}(P)} & W^{\text{elect}}>W^{\text{mag}}\\[2ex] \dfrac{2\omega W^{\text{mag}}}{\text{Re}(P)} & W^{\text{mag}}>W^{\text{elect}}\end{cases} \tag{20}$$

for this ideal antenna must be less than or equal to that for any other loss-free antenna fitting into the sphere $r=R$, since any field $r < R$ can only add to the energy storage. In (20) the energies W are those obtained by subtracting the radiation field from the total field. If the Q of an antenna is high, it can be interpreted as the reciprocal of the fractional bandwidth of the input impedance. If the Q is low the antenna has broadband potentialities.

Because of the orthogonality of the spherical wave functions, the total electric energy, magnetic energy, and power radiated is the sum of the corresponding quantities associated with each mode. The Q can therefore be found by treating the field of each spherical wave as if it existed on a "spherical waveguide" isolated from all other waves [2,3]. The energy and power formulas of transmission line theory apply to each spherical waveguide if a voltage, current, and characteristic impedance are defined for each TE_{mn} wave as

$$\begin{aligned} V_{mn}^{TE}&=\frac{A_{mn}}{\beta}\sqrt{\frac{4\pi n(n+1)(n+m)!}{\eta\epsilon_m(2n+1)(n-m)!}}\,F_n(\beta r)\\ I_{mn}^{TE}&=\frac{A_{mn}}{\beta}\sqrt{\frac{4\pi n(n+1)(n+m)!}{\eta\epsilon_m(2n+1)(n-m)!}}\,jF_n'(\beta r)\\ Z_{mn}^{TE}&=F_n(\beta r)/jF_n'(\beta r)\end{aligned} \tag{21}$$

and for each TM_{mn} wave as

$$\begin{aligned} V_{mn}^{TM}&=\frac{B_{mn}}{\beta}\sqrt{\frac{\eta 4\pi n(n+1)(n+m)!}{\epsilon_m(2n+1)(n-m)!}}\,jF_n'(\beta r)\\ I_{mn}^{TM}&=\frac{B_{mn}}{\beta}\sqrt{\frac{\eta 4\pi n(n+1)(n+m)!}{\epsilon_m(2n+1)(n-m)!}}\,F_n(\beta r)\\ Z_{mn}^{TM}&=jF_n'(\beta r)/F_n(\beta r).\end{aligned} \tag{22}$$

An antenna adjusted for maximum gain has equal excitation of TE and TM waves. It is therefore convenient to define modal quality factors

$$Q_n=\frac{2\omega W_{mn}^{\text{elect}}}{\text{Re}(P_{mn})}=\frac{2\omega W_{mn}^{\text{mag}}}{\text{Re}(P_{mn})}, \tag{23}$$

where the W_{mn} and P_{mn} are the sum of the energies and powers of both the TE_{mn} and TM_{mn} waves. Note that the Q_n's are independent of m since the characteristic impedances, (21) and (22), are independent of m. These are the same Q_n's defined by Chu for circularly polarized omnidirectional antennas [2]. Abstracting from Chu's work, one has

$$Q_n(\beta R)=\tfrac{1}{2}|F_n(\beta R)|^2\beta R X_n'(\beta R) \tag{24}$$

where R is the radius of the spherical antenna and

$$X_n(\beta R)=\text{Im}\,[Z_{mn}^{TM}(\beta R)]. \tag{25}$$

These Q_n are approximately one-half those calculated by Chu for linearly polarized omnidirectional antennas. A plot of some Q_n is given in figure 4.

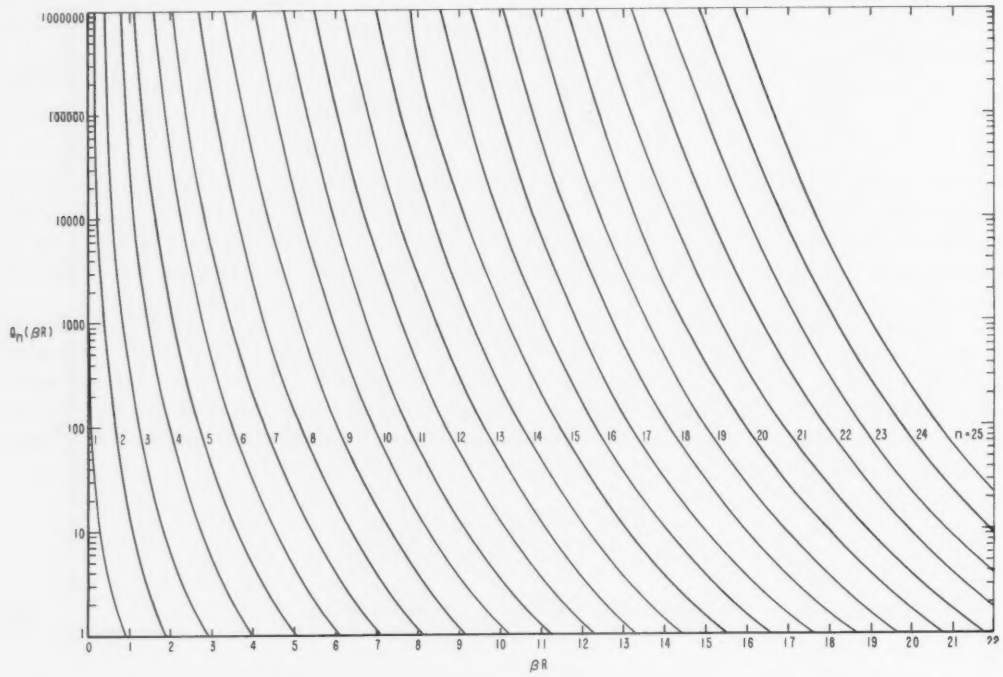


FIGURE 4. *Modal quality factors for wave functions of order* n.

The Q of an antenna having equal excitation of TE and TM waves is given by

$$Q=\frac{\sum_{m,n}\text{Re}(P_{mn})Q_n}{\sum_{m,n}\text{Re}(P_{mn})} \tag{26}$$

since it is merely necessary to add the energies and powers of the individual waves. For maximum gain all $P_{mn}=0$ except P_{1n}, $n\leq N$. It follows from (9), (11), (21), and (22) that

$$\begin{aligned}\text{Re}\,(P_{1n})&=\text{Re}\,(V_{1n}^{TE}I_{1n}^{TE*}+V_{1n}^{TM}I_{1n}^{TM*})\\&=\frac{4\pi}{\eta\beta^2}\frac{1}{2n+1}|a_n|^2\end{aligned} \tag{27}$$

so the formula for Q becomes

$$Q=\frac{\sum_{n=1}^{N}|a_n|^2\left(\frac{1}{2n+1}\right)Q_n}{\sum_{n=1}^{N}|a_n|^2\left(\frac{1}{2n+1}\right)}. \tag{28}$$

Finally, from (17) one has

$$Q=\frac{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2 Q_n(\beta R)}{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2} \tag{29}$$

as the quality factor for an ideal antenna of radius R, adjusted for maximum gain at a radius r, using wave functions of order $n \leq N$. For an antenna adjusted for maximum gain at infinity

$$Q \xrightarrow[\beta r \to \infty]{} \frac{\sum_{n=1}^{N}(2n+1)Q_n(\beta R)}{N^2+2N}. \tag{30}$$

Figure 5 shows the Q of antennas focused at infinity for several N. (The dashed lines represent

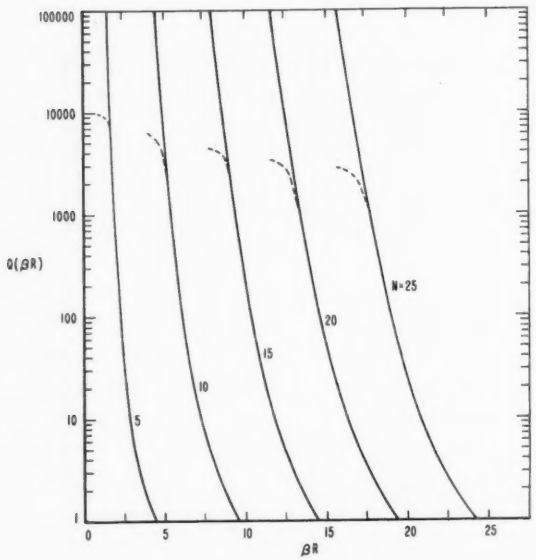


FIGURE 5. *Quality factors for ideal loss-free antennas adjusted for maximum gain at infinity, using wave functions of order* $n \leq N$ *(dashed lines show the effect of losses).*

the effect of antenna losses, considered below.) Calculations for antennas focused at other r show that the Q is substantially independent of the radius of focus. No appreciable change in Q occurs until r is almost equal to R, that is, the field point is almost at the antenna surface. When $r=R$, the Q of the antenna is close to the Q_n of the highest order wave present. When $\beta R > N$, all Q_n are of the order of unity or less, and the quality factor is

$$Q \leq 1 \qquad \beta R > N. \tag{31}$$

In this case the antenna is potentially broad band.

4. Effect on Antenna Losses

To obtain quantitative results for the effect of conduction losses on antenna performance an idealized model is again postulated. Intuitively one would expect the losses on a metal antenna to be smaller the more effectively the sphere is utilized. Therefore, for the ideal lossy antenna is postulated a spherical conductor of radius R excited by the magnetic sources

$$\mathbf{M}=\mathbf{E}\times\mathbf{n} \tag{32}$$

on its surface. So long as the sphere is a good conductor the source of (32) will generate the desired field [4]. If the conductivity of the sphere is poor, (32) can be modified to allow for a field internal to the conductor.

The above postulated ideal lossy antenna is particularly simple to analyze because the wave functions are orthogonal over its surface. The effect of the spherical conductor is that of a discontinuity in the characteristic impedance of each "spherical waveguide" at the radius R. The effect of the source **M** is that of a voltage source in series with each waveguide at the radius R. The waveguides are matched in each direction, so the equivalent circuit for each mode is a voltage source in series with the two characteristic impedances, $r<R$ and $r>R$. For $r<R$ the characteristic impedances for the various modes are

$$Z_{mn}^{TE}=\frac{\eta_c}{\eta}[F_n(kr)/jF_n'(kr)]^*\approx\eta_c/\eta$$
$$Z_{mn}^{TM}=\frac{\eta_c}{\eta}[jF_n'(kr)/F_n(kr)]^*\approx\eta_c/\eta \tag{33}$$

where k and η_c are the wave number and intrinsic impedance in the conductor,

$$k\approx(1-j)\sqrt{\frac{\omega\mu\sigma}{2}}\qquad \eta_c\approx(1+j)\sqrt{\frac{\omega\mu}{2\sigma}}. \tag{34}$$

The characteristic impedance $r<R$ is extremely small for good conductors, so $V_{mn}\approx 0$, $r<R$. The current I_{mn} must be continuous at $r=R$. Hence for any mode the ratio of power dissipated to power radiated is given by

$$\frac{P_{diss}}{P_{rad}}=\frac{|I_{mn}|^2\,\mathrm{Re}(Z_{mn}^-)}{|I_{mn}|^2\,\mathrm{Re}(Z_{mn}^+)}=\frac{\mathrm{Re}(\eta_c)}{\eta\mathrm{Re}(Z_{mn}^+)} \tag{35}$$

where the superscripts $+$ and $-$ refer to $r>R$ and $r<R$, respectively. Dissipation factors D_n are defined for the case of equal TE_{mn} and TM_{mn} excitation as

$$D_n=\frac{P_{diss}^{TE}+P_{diss}^{TM}}{P_{rad}^{TE}+P_{rad}^{TM}}=\frac{P_{diss}^{TE}}{2P_{rad}^{TE}}+\frac{P_{diss}^{TM}}{2P_{rad}^{TM}}$$
$$=\frac{\mathrm{Re}(\eta_c)}{2\eta}\left[\frac{1}{\mathrm{Re}(Z_{mn}^{TE})}+\frac{1}{\mathrm{Re}(Z_{zn}^{TM})}\right]. \tag{36}$$

The D_n are independent of m because the Z_{mn} are independent of m. Using (21) and (22) one has

$$D_n(\beta R)=\frac{\mathrm{Re}(\eta_c)}{2\eta}[|F_n'(\beta R)|^2+|F_n(\beta R)|^2]$$
$$=\frac{\mathrm{Re}(\eta_c)}{2\eta}[|u_n(\beta R)|^2-2] \tag{37}$$

where the $|u_n|^2$ are plotted in figure 1. Note that the D_n are essentially proportional to the $|u_n|^2$ when $\beta R<n$.

The dissipation factor for an antenna having equal excitation of TE and TM waves is defined as

$$D=\frac{P_{diss}}{P_{rad}}=\frac{\sum_{m,n}\mathrm{Re}(P_{mn})D_n}{\sum_{m,n}\mathrm{Re}(P_{mn})} \tag{38}$$

where P_{mn} is $P_{mn}^{TE}+P_{mn}^{TM}$. The second equality of (38) follows from the orthogonality property of the modes. For maximum gain all $P_{mn}=0$ except P_{1n}, $n\leq N$, which are given by (27). Using this and (17) one has

$$D=\frac{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2 D_n(\beta R)}{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2} \tag{39}$$

as the dissipation factor for an ideal lossy antenna of radius R, adjusted for maximum directive gain at a radius r, using wave functions of order $n\leq N$. For an antenna adjusted for maximum gain at infinity

$$D\xrightarrow[\beta r\to\infty]{}\frac{\sum_{n=1}^{N}(2n+1)D_n(\beta R)}{N^2+2N}. \tag{40}$$

Figure 6 shows the dissipation factor of antennas focused at infinity for several N. Calculations for antennas focused at other values of r show that D is essentially independent of the radius

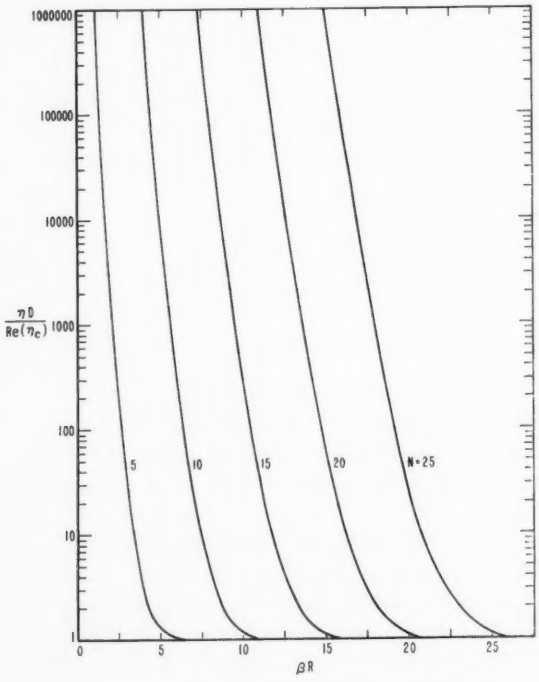


FIGURE 6. *Dissipation factors* (D) *for ideal lossy antennas adjusted for maximum gain at infinity, using wave functions of order* n≤N.

of focus, except when $r\approx R$. When $r=R$, the D of the antenna is approximately the D_n of the highest order wave present. If $\beta R>N$, all D_n are approximately $\mathrm{Re}(\eta_c)/\eta$, and the dissipation factor is

$$D\approx \mathrm{Re}\ (\eta_c)/\eta \tag{41}$$

and the antenna has very small loss. The efficiency of the antenna is

$$\%\ \text{efficiency}=\frac{100\,P_{\text{rad}}}{P_{\text{rad}}+P_{\text{diss}}}=\frac{100}{1+D}. \tag{42}$$

The antenna remains reasonably efficient until D is of the order of unity. The dissipation also changes the effective Q of the antenna, defined as

$$Q_{ett}=\begin{cases}\dfrac{2\omega W^{\text{elect}}}{P_{\text{rad}}+P_{\text{diss}}} & W^{\text{elect}}>W^{\text{mag}}\\[2ex] \dfrac{2\omega W^{\text{mag}}}{P_{\text{rad}}+P_{\text{diss}}} & W^{\text{mag}}>W^{\text{elect}}\end{cases} \tag{43}$$

Because of the orthogonality of energy and power, $P_{\text{rad}}+P_{\text{diss}}$ is simply the sum of the corresponding quantities for each mode, that is,

$$P_{\text{rad}}+P_{\text{diss}}=\sum_{m,n}(P_{mn})_{\text{rad}}(1+D_n). \tag{44}$$

Thus, for an antenna containing only the 1, n modes, $n\leq N$, one has instead of (28)

$$Q_{ett}=\frac{\sum_{n=1}^{N}|a_n|^2\left(\frac{1}{2n+1}\right)Q_n}{\sum_{n=1}^{N}|a_n|^2\left(\frac{1+D_n}{2n+1}\right)}. \tag{45}$$

Finally, if the antenna is adjusted for maximum directive gain, the a_n are given by (17) and

$$Q_{ett}=\frac{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2Q_n(\beta R)}{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2[1+D_n(\beta R)]}. \tag{46}$$

If the antenna is adjusted for maximum gain at infinity this becomes

$$Q_{ett}\xrightarrow[\beta r\to\infty]{}\frac{\sum_{n=1}^{N}(2n+1)Q_n(\beta R)}{\sum_{n=1}^{N}(2n+1)[1+D_n(\beta R)]}. \tag{47}$$

The dashed lines in figure 5 show the effective Q for various N assuming $\text{Re}(\eta_c)/2\eta=10^{-4}$, which corresponds to good conductors in the vicinity of 10,000 Mc. Further calculations show that the effective Q is essentially independent of the radius of focus, just as in the loss-free case. For large antennas, the maximum effective Q is of the order of that for a good spherical resonator constructed of the same metal.

While the directive gain of an antenna is unaffected by dissipation (assuming that the current distribution is unchanged), the overall gain

$$g(\beta r)=\frac{4\pi r^2\,\text{Re}\,(S_r)_{\max}}{P_{\text{rad}}+P_{\text{d'ss}}} \tag{48}$$

is affected. This is the gain usually of primary interest for antenna evaluation. One can quite simply go back and maximize g, since the P_{diss} of each mode is related to the P_{rad} of each mode by (35). The only difference in the equations for G and those for g is that the factor $1/(2n+1)$ is replaced by $(1+D_n)/(2n+1)$. Hence, (16) becomes [4]

$$g_{\max}=\frac{1}{4}\sum_{n=1}^{N}\frac{(2n+1)}{1+D_n(\beta R)}\,|u_n(\beta r)|^2 \tag{49}$$

[4] This procedure is slightly in error since dissipation factors for TE and TM modes alone are not quite equal. The correction is small, however, until the dissipation factors become quite large.

and the wave amplitudes, related previously by (17), are related by

$$\frac{|a_n|}{|a_i|}=\frac{(1+D_i)(2n+1)|u_n|}{(1+D_n)(2i+1)|u_i|} \tag{50}$$

Note that the maximum gain is now a function of R (antenna size) as well as r (field point). If the antenna is focused at infinity (49) reduces to

$$g_{max}\xrightarrow[\beta r\to\infty]{}\sum_{n=1}^{N}\frac{2n+1}{1+D_n(\beta R)}. \tag{51}$$

So long as $D_N<1$ the overall gain g is substantially equal to the directive gain G. If $D_N>1$ the maximum gain depends upon the surface resistance of the metal, $\mathrm{Re}(\eta_c)$. However, the $|u_n|^2$ functions, which enter into D_n according to (37), rise very rapidly. For good conductors the "cut-off" of the summation of (46) occurs at approximately the same value of n as that for which the Q_n level off.

When the overall gain instead of the directive gain is maximized it affects the dissipation and quality factors. The dissipation factor becomes

$$D=\frac{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2D_n(\beta R)/[1+D_n(\beta R)]^2}{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2/[1+D_n(\beta R)]^2} \tag{52}$$

instead of (39). The principal difference between (39) and (52) is that in the latter case D levels off when the D_n becomes greater than unity. This can be thought of as due to the non-utilization of modes which are highly dissipative. The quality factor becomes

$$Q_{eff}=\frac{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2Q_n(\beta R)/[1+D_n(\beta R)]^2}{\sum_{n=1}^{N}(2n+1)|u_n(\beta r)|^2/[1+D_n(\beta R)]} \tag{53}$$

instead of (46). A plot of (53) would again give curves similar to the dashed lines of figure 5.

5. Discussion

To relate the analysis to practical antenna systems, define the radius R of an antenna system to be the radius of the smallest sphere that can contain it. The Q of the ideal loss-free antenna must then be less than or equal to the Q of any other loss-free antenna of radius R, since fields $r<R$ can only add to energy storage. In other words, the Q of figure 5 is a lower bound to the Q of an arbitrary loss-free antenna of radius R. It would be nice if one could also prove that the dissipation factor D of the ideal lossy antenna were a lower bound to the D of an arbitrary antenna of the same material and radius. The author has not been able to prove this. However, it will be assumed that the D of the ideal lossy antenna is of the same order of magnitude as for other antennas of the same material and radius. Calculations of the D for some practical antennas support this assumption.

It is evident from the foregoing analysis that a marked change in the behavior of an antenna of radius R occurs when wave functions of order $n>\beta R$ are present in its field. In the loss-free case the Q is large. In addition to this, in the lossy case the dissipation is large. In both cases the near-field is characterized by extremely large field intensities. The *normal gain* of an antenna is defined to be the maximum gain obtainable using wave functions of order $n\leq N=\beta R$. Hence, from (19), the normal gain of an antenna of radius R is

$$G_{\text{norm}}=(\beta R)^2+2(\beta R) \tag{54}$$

in the radiation zone. Systems having larger gain than this are called *supergain antennas.* For large βR the gain of a uniformly illuminated aperture of radius R is equal to the above defined normal gain [5]. Therefore, one cannot obtain a gain higher than that of the uniformly illuminated aperture without resorting to a supergain antenna.

It is evident from figure 3 that the maximum near-zone gain of an antenna using wave functions $n \leq N = \beta R$ is essentially the same as the maximum far-zone gain. Hence (54) also defines the normal near-zone gain for all practical purposes. A uniformly illuminated and "focused" aperture (phase adjusted so that all elements contribute in-phase at some distance r) has a near-zone gain approximately equal to the far-zone gain of the "unfocused" aperture (uniform phase). Thus, a near-zone gain greater than that obtainable from a focused uniformly illuminated aperture cannot be obtained without resorting to a supergain antenna.

Having precisely defined the term "supergain," one can now consider the question of how much supergaining is possible. If a particular Q is chosen, the possible increase in gain over normal gain can be readily calculated. For example, if $Q = 10^6$ is taken, the decibel increase in gain over normal gain is as shown in figure 7. Note that for small R, substantial increases in gain can be achieved, but for large R the increase becomes insignificant. The curve of figure 7 is relatively insensitive to the particular choice of Q, so long as it is high. This is evident from the rapid rise of the curves of figure 5. The choice $Q = 10^6$ represents sort of an upper limit to practically significant Q's, since the bandwidth becomes absurdly narrow for higher Q's. Also, $Q = 10^6$ represents the approximate upper limit for antennas constructed of metal, due to dissipation.

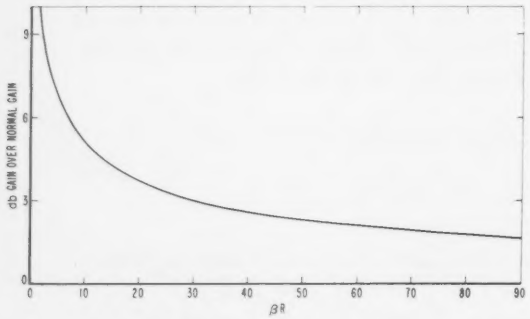


FIGURE 7. *Maximum increase in gain over normal gain obtainable by supergaining if a* Q *of* 10^6 *is allowed.*

It is evident from figure 5 that the amount of supergaining possible in large antennas is very small. Hence, for practical purposes, the uniformly illuminated aperture gives optimum gain. For small antennas, however, a significant increase over normal gain is possible. Perhaps the most common example of a small supergain antenna is the short dipole. The problems of narrow bandwidth and high losses associated with this antenna have been thoroughly treated, since it is one of the few antennas that can be used at very low frequencies.

6. References

[1] R. F. Harrington, On the gain and beamwidth of directional antennas, IRE Trans. **AP–6,** 219 (1958).
[2] L. J. Chu, Physical limitations of omnidirectional antennas, J. Appl. Phys. **19,** 1163 (1948).
[3] N. Marcuvitz, Waveguide handbook, pp. 96–100 (McGraw-Hill Book Co., New York, N.Y., 1951).
[4] R. F. Harrington, Introduction to electromagnetic engineering, pp. 167–177 (McGraw-Hill Book Co., New York, N.Y., 1958).
[5] S. Silver, Microwave antenna theory and design, p. 177 (McGraw-Hill Book Co., New York, N.Y., 1949).

BOULDER, COLO.

(Paper 64D1–31)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Surface-Wave Resonance Effect in a Reactive Cylindrical Structure Excited by an Axial Line Source [1]

A. L. Cullen [2]

(July 27, 1959)

It is shown that a purely reactive cylinder excited by a neighboring line source can, under suitable conditions, give rise to a radiation pattern closely approximating the function $\cos n\theta$.

In a numerical example, a cylinder of three transverse electro-magnetic (T.E.M.) wavelengths circumference has a surface reactance chosen to emphasize the term $\cos 6\theta$ in the Fourier series of the resultant radiation pattern. It is shown that only 1.1 percent of the total power delivered to the line source is radiated in unwanted modes.

It is also shown that the position of the line source does not affect this result to first order provided that $k(b-a)\ll 1$, where $b-a$ is the distance of the line source from the cylindrical surface.

1. Introduction

In a recent paper the natural (exponentially decaying) oscillations of a highly reactive cylinder supporting a circulating surface wave were discussed [1].[3] It was shown that for cylinders of a few wavelengths circumference, Q-factors of the order of several hundred are possible, if the circumferential phase velocity of the surface waves is appreciably less than the velocity of light.

The purpose of the present paper is to examine the corresponding forced oscillations, taking a line source as the means of excitation.

The problem of a line source in the vicinity of a cylinder with a reactive surface impedance has already been solved [2,3] by Wait and Conda, but the reactance values considered in their numerical studies were not such as to exhibit strong resonance effects. Also, they considered a magnetic current line source, whereas for the surface waves discussed in [1] an electric current line source is appropriate. This is a trivial matter, however, since the principle of duality can be employed to pass from one case to the other.

2. Formulation of the Problem

The practical interest in strong resonance phenomena lies in the possibility that a radiation pattern of the form $a_n \cos n\theta$ might be set up very simply by employing the combination of a line source and a reactive cylindrical surface. Inspection of the general formula shows at once that terms of the form $a_m \cos m\theta$ will always be present, but when the surface reactance is chosen to establish resonance with the reactive part of the field impedance of, say, the nth mode, we might expect to find $a_n \gg a_m$, $m \neq n$.

If we adopt the mean-square-deviation from $a_n \cos n\theta$ as the measure of unwanted radiation, it is obvious that we must maximize the following expression:

$$\eta_n = \frac{\frac{1}{2}|a_n|^2}{|a_0|^2 + \frac{1}{2}\sum_{m=1}^{\infty}|a_m|^2} \qquad (n \neq 0).$$

[1] The research reported in this paper was sponsored by the U.S. Air Force Cambridge Research Center under contract CSO and A 58-40.
[2] Guest researcher from the Department of Electrical Engineering, University of Sheffield, Sheffield, England.
[3] Figures in brackets indicate the literature references at the end of this paper.

The quantity η_n may be called the "excitation efficiency" or "pattern purity" for the nth mode.

The excitation efficiency defined in this way has a very close affinity to the launching efficiency for surface waves of a line source over a plane reactive surface [4]. In this case, the launching efficiency of the surface wave depends markedly on the height h of the source above the reactive surface, the principal effect being a great reduction in radiated power for a certain value of h.

It seems reasonable to expect that a similar effect might occur in the cylindrical case if we take $\frac{1}{2}|a_n^2|$ as analogous to the surface-wave power and

$$|a_o^2|+\frac{1}{2}\sum_{m=1}^{n-1}|a_m^2|+\frac{1}{2}\sum_{n+1}^{\infty}|a_m^2|$$

as analogous to the radiated power in the plane surface wave excitation problem. With this analogy in mind, one might look for an optimum position of the line source which would give the closest approximation, in the least-mean-squares sense, to a $\cos n\,\theta$ radiation pattern.

3. Current Sheet Excitation

Although the necessary formulas have already been given by Wait and Conda, we shall sketch a derivation for completeness. The opportunity will be taken to adapt the point of view employed in [4] to the cylindrical case. That is to say, the line source is synthesized from cylindrical current sheets rather than from point dipoles as in Wait and Conda's work. Electric current sources will be considered rather than magnetic current sources for the reason mentioned earlier.

Referring to figure 1, a sinusoidally distributed cylindrical current sheet of radius b coaxial with the cylinder of radius a, $(a \leqq b)$ is first considered. The current is assumed to flow parallel to the axis, and the current density is written

$$K_z=K_n \cos n\,\theta \tag{1}$$

where n must clearly be an integer if K_z is to be a single-valued function of θ.

The following equations for the fields can then be written down

$a \leqq r \leqq b$

$$E_z'=[AJ_n(kr)+BH_n(kr)]\cos n\,\theta \tag{2}$$

$$H_\theta'=\frac{1}{j\zeta}[AJ_n'(kr)+BH_n'(kr)]\cos n\,\theta \tag{3}$$

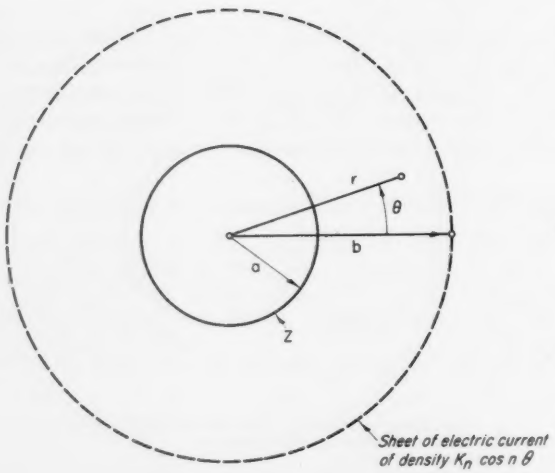


FIGURE 1. *Surface-current sheet excitation for cylinder having finite surface impedance.*

$b<r$

$$E_z = CH_n(kr)\cos n\theta \tag{4}$$

$$H_\theta = \frac{C}{j\zeta} H_n'(kr)\cos n\theta \tag{5}$$

In the above equations $H_n^{(2)}(kr)$ has been written as $H_n(kr)$ to simplify the notation. No confusion should result since the function $H_n^{(1)}(kr)$ is not employed in this paper. The quantity ζ is the intrinsic impedance of free space $(\mu_0/\epsilon_0)^{\frac{1}{2}}$.

The constants A, B, and C are to be evaluated from the conditions

(i) $E_z' = ZH_\theta'$ at $r=a$

where Z is the surface impedance looking *into* the cylinder.

(ii) $E_z' = E_z$ at $r=b$

(iii) $H_\theta' - H_\theta = K_z$ at $r=b$.

After straightforward algebra, making use of the Wronskian to simplify the term $J_nH_n' - H_nJ_n'$ which arises, we arrive at the following expression for E_z for $r>b$,

$$E_z = \frac{\pi b}{2}\omega\mu_0 K_n[J_n(kb)+\alpha_n H_n(kb)]H_n(kr)\cos n\theta \tag{6}$$

where

$$\alpha_n = -\frac{j\zeta J_n(ka) - ZJ_n'(ka)}{j\zeta H_n(ka) - ZH_n'(ka)}. \tag{7}$$

Before proceeding to the next step of synthesizing the line source from current sheet sources of the form described by eq (1), we shall examine (7) which plays a role roughly similar to the reflection coefficient $(u+k_1)/(u-k_1)$ of the plane surface wave excitation case [4]. However, the infinite reflection coefficient found for plane surface waves is not to be expected with circulating cylindrical surface waves, for these waves are inherently leaky, and so damping is present even when the impedance Z is purely reactive. In fact we shall show that $|\alpha_n|$ cannot exceed unity.

α_n will first be written in terms of J_n and Y_n, putting $Z=jX$.

$$\alpha_n = -\frac{\left[J_n(ka) - \frac{X}{\zeta}J_n'(ka)\right]}{\left[J_n(ka) - \frac{X}{\zeta}J_n'(ka)\right] - \mathrm{j}\left[Y_n(ka) - \frac{X}{\zeta}Y_n'(ka)\right]}. \tag{8}$$

For $n>ka$ pronounced resonance effects occur. Y_n and Y_n' are in general considerably larger than J_n and J_n' so that α_n tends to be small unless $Y_n(ka) - (X/\zeta)\ Y_n'\ (ka)$ is zero.

For $n<ka$, however, the resonance effect is very weak.

The condition for resonance can be written

$$\frac{X}{\zeta} = \frac{Y_n(ka)}{Y_n'(ka)}. \tag{9}$$

When (9) is satisfied

$$\alpha_n = -1. \tag{10}$$

If (9) is not satisfied it follows from eq (8) that

$$|\alpha_n| < 1. \tag{11}$$

Thus, we have proved the statement made earlier that the quantity α_n is never infinite, and in fact cannot exceed unity.

A closer analogy with reflection coefficient is the complex ratio of outgoing to incoming cylindrical waves in the region $a<r<b$. This is equal to $1+2\alpha_m$, and the reflection coefficient in this sense can never exceed unity in magnitude.

As a numerical example of the way the coefficients α_n depend on n, $|\alpha_n|^2$ has been calculated as a function of n for the case in which X/ζ is chosen to satisfy the resonance condition, eq (9), for $n=6$, taking $ka=3$. The appropriate value of X/ζ is -0.605. The results are given in table 1.

TABLE 1

n	$\lvert\alpha_n^2\rvert$
0	0.869
1	.054
2	.970
3	.641
4	.151
5	.018
6	1.000
7	1.4×10^{-6}

The very rapid falling-off of $|\alpha_n|^2$ with n for n larger than the resonant value is indicated by the last line of this table.

The sharpness of the resonance of $|\alpha_6|^2$ as X/ζ is varied is shown in figure 2.

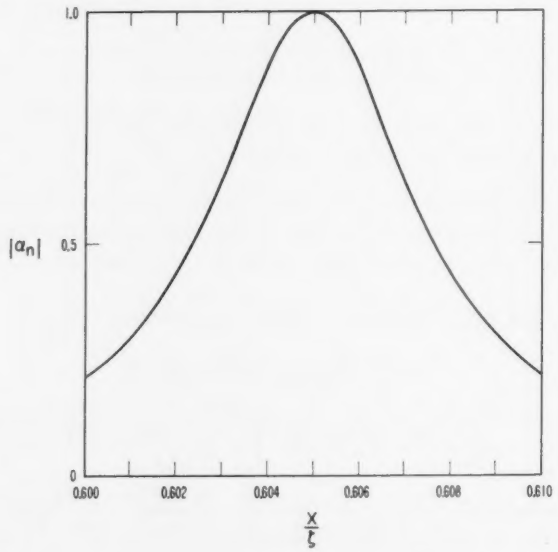


FIGURE 2. *Variation of $|\alpha_n|$ with X/ζ, showing the effect of surface-wave resonance.*

If we define a Q-factor for the variation of $|\alpha_n|^2$ with X through the equation

$$|a_n|^2 \cong \frac{1}{1+\left(Q_x \frac{\delta X}{X_0}\right)^2} \tag{12}$$

the approximate Q-factor deduced from figure 2 is $Q_x=230$.

This is considerably lower than the Q-factor deduced in [1], but the conditions are somewhat different.

The difference lies in the way the reactances involved depend on the frequency. In [1] it is assumed that the capacitive surface reactance X is inversely proportional to the frequency, while the inductively reactive part of the field impedance increases with frequency, in accordance with field theory, more rapidly than in direct proportion to the frequency.

The Q-factor Q_x defined by (12) would only agree with the Q-factor of [1] if the reactive part of the field impedance varied in direct proportion to the frequency.

Hence, the Q-factor of [1] will necessarily exceed the rather artificial Q_x introduced here.

4. Line Source Excitation

Suppose we have a linear current filament I at $r=b$, $\theta=0$. This can be represented as a Fourier series thus

$$\frac{I\delta(\theta)}{b}=\sum_{m=0}^{\infty}K_m \cos m\theta. \tag{13}$$

The coefficients K_m represent the amplitudes of cosinusoidally distributed surface current densities, and are found in the usual way to be

$$K_m=\epsilon_m\frac{I}{2\pi b} \tag{14}$$

where $\epsilon_0=1$, $\epsilon_m=2$, $m\neq 0$.

The field associated with each K_m can be found from eq (6), and in this way we find the formula for current filament excitation

$$E_z=\frac{\omega\mu_0 I}{4}\sum_{m=0}^{\infty}\epsilon_m[J_m(kb)+\alpha_m H_m(kb)]H_m(kr) \cos m\theta. \tag{15}$$

This formula makes possible far-field calculations, and for this purpose $H_m(kr)$ can be replaced by its asymptotic value.

$$H_m(kr)\cong\left(\frac{2}{\pi kr}\right)^{\frac{1}{2}}\exp\left[-j\left(kr-\frac{1}{2}m\pi-\frac{1}{4}\pi\right)\right]. \tag{16}$$

Equation (15) is the dual of the formula derived by Wait [3] and is identical apart from minor differences of notation if the following transformations are made: $E_z\rightarrow H_z$, $\mu_0\rightarrow\epsilon_0$, and $I\rightarrow K$ (symbol used to denote magnetic current). It may also be derived from Wait's result [5] for the fields of an electric dipole in the presence of a cylinder with impedance boundary conditions.

Equation (15) enables us to calculate the pattern purity or excitation efficiency according to the definition given in the introduction to this paper. The asymptotic value of $H_m(kr)$ given by (16) has the magnitude $(2/\pi kr)^{1/2}$ for all values of m, and therefore we may take as our a_m's the quantities $\epsilon_m[J_m(kb)+\alpha_m H_m(kb)]$.

Thus we find

$$\eta_n=\frac{2|J_n(kb)+\alpha_n H_n(kb)|^2}{\sum_{m=0}^{\infty}\epsilon_m|J_m(kb)+\alpha_m H_m(kb)|^2} \qquad (n\neq 0). \tag{17}$$

The special case $b=a$ is worthy of attention. In this case the term in square brackets in eq (15) can be simplified by making use of the Wronskian again, to read

$$\frac{2j}{\pi ka}\left[\frac{\frac{X}{\zeta}}{H_n(ka)-\frac{X}{\zeta}H_n'(ka)}\right] \tag{18}$$

and eq (15) then becomes

$$E_z=\frac{j\zeta I}{2\pi a}\sum_{m=0}^{\infty}\epsilon_m\left[\frac{\frac{X}{\zeta}}{H_m(ka)-\frac{X}{\zeta}H'_m(ka)}\right]H_m(kr)\cos m\theta. \tag{19}$$

The eq (17) for excitation efficiency can be specialised for $b=a$ to read

$$\eta_n=\frac{2\left|\frac{1}{H_n(ka)-\frac{X}{\zeta}H'_n(ka)}\right|^2}{\sum_{m=0}^{\infty}\epsilon_m\left|\frac{1}{H_m(ka)-\frac{X}{\zeta}H'_m(k_a)}\right|^2}. \tag{20}$$

Let us now consider (17) for the case $k(b-a)<<1$.
We can write

$$J_m(kb)+\alpha_m H_m(kb)\cong\{J_m(ka)+\alpha_m H_m(ka)\}+\{J'_m(ka)+\alpha_m H'_m(ka)\}k(b-a)$$

if terms of order $k^2(b-a)^2$ are neglected. We can use (18) to replace the first term on the right-hand side of this equation, and by a similar simplification of the second term we find

$$J_m(kb)+\alpha_m H_m(kb)=\frac{2j}{\pi ka}\left[\frac{\frac{X}{\zeta}}{H_n(ka)-\frac{X}{\zeta}H'_n(ka)}\right]\times\left\{1+\frac{\zeta}{X}k(b-a)\right\}. \tag{21}$$

The last factor in (21) is independent of m, and so to a first approximation we find that all modes are multiplied by the same factor. Thus, although a small finite value of $k(b-a)$ will alter the total power supplied by the source for a given current, and so will alter the radiation resistance of the line source, it will not affect the distribution of power among the various modes, and so will not change the excitation efficiency.

Thus it follows that eq (20) is valid for $b\neq a$, provided that $k(b-a)<<1$.

5. Example of Excitation Efficiency Calculation

We now consider the excitation efficiency of the mode $n=6$ by a line source near the cylinder, i.e., $k(b-a)<<1$, for the case $ka=3$. The surface reactance is again chosen to establish resonance for the mode $n=6$. Remembering that excitation efficiency, or pattern purity, is defined as power in the wanted mode divided by the total radiated power, i.e., in symbols

$$\eta_n=\frac{P_n}{\sum_{m=1}^{\infty}P_m}$$

it is instructive to tabulate the P_m's in order to find how the unwanted radiation is distributed amongst the various modes. This information is given in table 2.

TABLE 2

m	P_m
0	4.02
1	8.00
2	7.85
3	7.36
4	6.74
5	6.78
6	3580.
7	0.08

It appears that very little power is radiated in modes for which $m>6$, and that for $m<6$, the power is fairly equally divided between the various modes.

Only 1.1 percent of the total power goes into unwanted modes, and the excitation efficiency has the very high value of 98.9 percent.

6. Conclusions

The existence of strong resonance phenomena associated with surface waves circulating round a reactive cylinder excited by a line source has been established theoretically. In a particular case, 98.9 percent of the total input power is radiated in the resonant mode, i.e., the excitation efficiency, or pattern purity is 98.9 percent.

It is shown that to first order in $k(b-a)$, the distance of the line source from the surface of the cylinder does not affect the excitation efficiency, though the radiation resistance of the source will decrease as $k(b-a)$ increases.

This paper was written during a short visit to the Boulder Laboratories of the National Bureau of Standards. The author expresses his gratitude for this most stimulating opportunity, particularly to James R. Wait of these laboratories through whom the visit was arranged.

7. References

[1] M. F. Bracey, A. L. Cullen, E. F. F. Gillespie, and J. A. Staniforth, Surface wave research in Sheffield, Proc. of URSI Symposium on Electromag. Wave Theory (1959).

[2] J. R. Wait and Alyce M. Conda, Radiation from slots on dielectric clad and corrugated cylinders, J. Research NBS, 59, 307 (1957) RP2802.

[3] J. R. Wait, Electromagnetic radiation from cylindrical structures, p. 135 (Pergamon Press, New York, N.Y., 1959).

[4] A. L. Cullen, The excitation of plane surface waves, Proc. Inst. Elec. Engrs. Part IV, 101, 225 (1954).

[5] J. R. Wait, Radiation from an electric dipole in the presence of a corrugated cylinder, Appl. Sci. Research, Sec. B, 6, 117 (1956).

Boulder, Colo. (Paper 64D1–32)

Journal of Research of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 1, January–February 1960

Basic Experimental Studies of the Magnetic Field from Electromagnetic Sources Immersed in a Semi-Infinite Conducting Medium

Martin B. Kraichman [1]

(July 20, 1959)

Using electromagnetic sources, consisting of various dipoles and loops immersed in a concentrated sodium chloride solution, measurements were made verifying the magnetic field propagation equations in air, derived previously by several authors. The receiver was farther away from the source than a wavelength in the conducting medium, but much closer than a wavelength in air.

An expression is derived giving the value of the magnetic field in air due to a rectangular loop with a horizontal axis by assuming the loop to consist of two electric dipoles corresponding to the horizontal members. Experimental data verifying this expression are presented.

Also, using submerged electric dipoles, measurements were made of the magnetic field in air which show that the field is determined solely by the current in the horizontal radiating wires of the dipoles.

1. Introduction

The purpose of the experimental work reported herein is to afford a better understanding of the electromagnetic propagation phenomena involved in radiation from sources immersed in a conducting medium.

Theoretical treatments of the propagation problem have been given by several authors, Tai [1],[2] Moore [2], Von Aulock [3], Wait [4], Wait and Campbell [5], Lien [6], and Baños and Wesley [7]. The important propagation path was shown to occur mainly through the air above the interface. The field along the direct path through the conducting medium, from source to receiver, suffered an exponential attenuation with the distance and was, therefore, negligible for all but very short distances.

2. Scope of Work

The experimental work undertaken at the U.S. Naval Ordnance Laboratory (NOL) consisted in checking the magnetic field propagation equations in air for the horizontal range, where the receiver is farther away from the source than a wavelength in the conducting medium, but much closer than a wavelength in air. Also, experiments were performed to determine the localization of the field source; i.e., whether the currents in the dipole could be considered as the sole contributors or whether currents in the conducting medium must also be considered in any discussion of the field.

3. Theory

In attempting to verify the propagation equations, it was decided to measure the magnetic field components using a shielded loop. This avoids grounding difficulties inherent in the use of an electric antenna. Also, because of the continuity of the magnetic field between air and salt water, measurements made at the surface in air will be valid just below the surface in the conducting medium.

The magnetic field propagation equations for a submerged dipole are given below in cylindrical coordinates using the notation of Baños and Wesley [7]. The cylindrical coordinate system and the position of the dipole are shown in figure 1. It is understood that the time factor $e^{-i\omega t}$ is to be applied to all field components.

for $K_2\rho \ll 1 < |K_1\rho|, K_2 z \ll 1$, and $\frac{z}{\rho} < 1$,

$$H\rho_2 = \frac{iIl \sin\phi}{\pi K_1 \rho^3} e^{iK_1 h} \tag{1}$$

$$H\phi_2 = \frac{-iIl \cos\phi}{2\pi K_1 \rho^3} e^{iK_1 h} \tag{2}$$

$$Hz_2 = \frac{-iIl \sin\phi}{2\pi K_1 \rho^3} \left(\frac{3}{iK_1\rho}\right)(1 - iK_1 z) e^{iK_1 h}, \tag{3}$$

[1] U.S. Naval Ordnance Laboratory, White Oak, Silver Spring, Md.

[2] Figures in brackets indicate the literature references at the end of this paper.

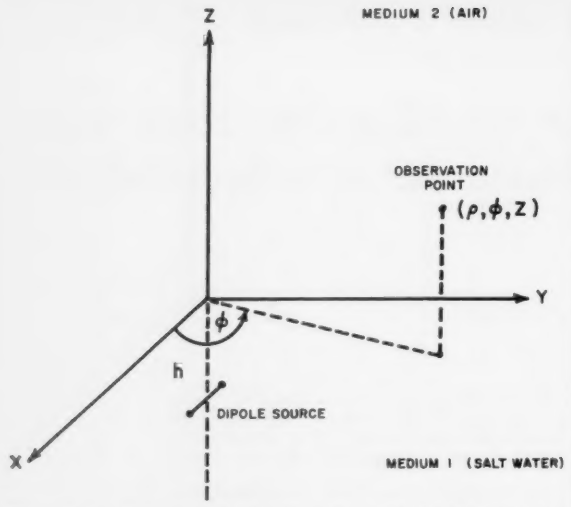


FIGURE 1. *Submerged dipole referred to cylindrical coordinate system.*

where

ω = angular frequency,
μ_0 = permeability of free space,
ϵ_0 = permittivity of free space,
σ = conductivity of medium 1 (salt solution),
K_1 = propagation constant for plane homogeneous waves in medium 1, $K_1^2 \cong i\omega\mu_0\sigma$,
K_2 = propagation constant for plane homogeneous waves in medium 2 (air), $K_2^2 \cong \omega\mu_0\epsilon_0$,
I = dipole current,
l = dipole length,

$\left.\begin{matrix}\rho\\ \phi\\ z\end{matrix}\right\}$ = cylindrical coordinates,

h = depth of dipole source.

If a traverse is made parallel to the x-axis with the radiating dipole positioned along the axis, a null is found for Hy_2. For each traverse the null lies along a line making an angle of approximately 35° with the x-axis. This may be derived from the propagation equations as follows:

$$Hy = H\rho \sin\phi + H\phi \cos\phi. \quad \text{For } Hy = 0,$$

$$H\rho \sin\phi = -H\phi \cos\phi.$$

Using expressions for $H\rho$ and $H\phi$ from eq (1) and (2),

$$\frac{iIl \sin^2\phi}{\pi K_1 \rho^3} e^{iK_1 h} = \frac{iIl \cos^2\phi}{2\pi K_1 \rho^3} e^{iK_1 h}$$

or

$$\tan^2\phi = \frac{1}{2}.$$

Thus

$$\phi = \tan^{-1}\frac{1}{\sqrt{2}} \cong 35°16'. \tag{4}$$

In order to verify the absolute value of the field given by eq (1), the expression for the radial component must be modified for calculation in the following way.

If $K_1 = (i\omega\mu_0\sigma)^{\frac{1}{2}} = i^{\frac{1}{2}}\sqrt{2}\,\beta$, where $\beta = \left(\frac{\omega\mu_0\sigma}{2}\right)^{\frac{1}{2}}$,

then

$$H\rho_2 = \frac{i^{\frac{1}{2}} Il \sin\phi}{\pi\sqrt{2}\,\beta\rho^3} e^{i^{\frac{3}{2}}\sqrt{2}\beta h}.$$

Also since $i^{\frac{1}{2}} = e^{i\frac{\pi}{4}}$ and $i^{\frac{3}{2}} = -\frac{1}{\sqrt{2}}(1-i)$ we get that

$$H\rho_2 = \left[\frac{Il}{\pi\sqrt{2}\,\beta}\frac{\sin\phi}{\rho^3} e^{-\beta h}\right] e^{i\left(\beta h + \frac{\pi}{4}\right)}. \tag{5}$$

The bracketed quantity in eq (5) gives the modulus of the radial magnetic field in air.

Equation (3) may be modified in a similar manner to yield an expression for the vertical magnetic field component.

$$Hz_2 = \left[\frac{3}{4}\frac{Il \sin\phi}{\pi\beta^2\rho^4} e^{-\beta h}\right] e^{i\left(\beta h + \frac{\pi}{2}\right)}$$

$$-\left[\frac{3\sqrt{2}Il \sin\phi}{4} \frac{z}{\pi\beta\rho^4} e^{-\beta h}\right] e^{i\left(\beta h + \frac{5}{4}\pi\right)}. \tag{6}$$

If

$$r_1 = \frac{3}{4}\frac{Il \sin\phi}{\pi\beta^3\rho^4} e^{-\beta h}$$

and

$$r_2 = \frac{-3\sqrt{2}\,Il \sin\phi}{4} \frac{z}{\pi\beta\rho^4} e^{-\beta h}$$

then

$$Hz_2 = [r_1 e^{i\frac{\pi}{2}} + r_2 e^{i\frac{5}{4}\pi}] e^{i\beta h}$$

$$= \left[\frac{-r_2}{\sqrt{2}} + i\left(r_1 - \frac{r_2}{\sqrt{2}}\right)\right] e^{i\beta h}$$

$$Hz_2 = \left[\frac{r_2^2}{2} + \left(r_1 - \frac{r_2}{\sqrt{2}}\right)^2\right]^{\frac{1}{2}} e^{i\beta h + \theta} \tag{7}$$

where

$$\theta = \tan^{-1}\frac{r_2 - \sqrt{2}r_1}{r_2}. \tag{8}$$

The bracketed quantity in eq (7) gives the modulus of the vertical component of the magnetic field in air.

In calculating the magnetic field due to a rectangular loop with a horizontal axis, the loop may be considered as being composed of horizontal and vertical electric dipoles of length corresponding to the horizontal and vertical members of the loop, respectively.

It has been shown by Von Aulock [3] that the magnetic field due to a vertical dipole attenuates exponentially with the distance from the surface above the dipole to the point of observation in the conducting medium. This results in a negligible contribution for any reasonable distance of observation from the dipole. Furthermore, the boundary conditions require that the magnetic field be continuous across the surface of the conducting medium. Thus, if measurements are made above the surface of the conducting medium. there should be no contribution to the field in air due to a vertical dipole. The field due to the rectangular loop may then be calculated on the basis of two horizontal members through which the excitation current is oppositely directed.

Considering the radial field, we get from eq (5)

$$H\rho_2=\left[\frac{Il}{\pi\sqrt{2}\beta}\frac{\sin\phi}{\rho^3}e^{-\beta h}\right]e^{i\left(\beta h+\frac{\pi}{4}\right)} \tag{9}$$

$$H'\rho_2=\left[\frac{Il}{\pi\sqrt{2}\beta}\frac{\sin\phi}{\rho^3}e^{-\beta h'}\right]e^{i\left(\beta h'+\frac{\pi}{4}\right)}. \tag{10}$$

Combining eq (9) and (10)

$$H\rho_2+H'\rho_2=\frac{Il\sin\phi}{\pi\sqrt{2}\beta\rho^3}e^{i\frac{\pi}{4}}\left[e^{-\beta h+i\beta h}-e^{-\beta h'+i\beta h'}\right]$$

or

$$H\rho_2+H'\rho_2$$

$$=\frac{Il\sin\phi}{\pi\sqrt{2}\beta\rho^3}\left[e^{-2\beta h}+e^{-2\beta h'}-2e^{-\beta(h+h')}\cos\beta\{(h-h')\}\right]^{\frac{1}{2}}e^{i\left(\theta+\frac{\pi}{4}\right)} \tag{11}$$

where

$$\theta=\tan^{-1}\frac{e^{-\beta h}\sin\beta h-e^{-\beta h'}\sin\beta h'}{e^{-\beta h}\cos\beta h-e^{-\beta h'}\cos\beta h'}. \tag{12}$$

The expression for the field of a rectangular loop given by eq (11) agrees with the integrated expression given by Von Aulock [3] for the field of a small coil with a horizontal axis.

4. Experimental Facilities

Rather than attempt a full scale verification of the propagation equations at sea, a reduced scale technique was adopted. This was done for a number of reasons. First, the cost of such an operation would be reduced considerably. Second, a land facility would permit wider exploration of experimental variables under more controlled conditions. Third, a facility on the base could readily be used for future investigations.

After balancing cost factors against convenience, the facility constructed consisted of an excavation 20 by 40 ft at the surface, 8-ft deep, with sides sloping at an angle of 45°. The sides and bottom were waterproofed with tar-impregnated felt.

A frequency of approximately 300 kc was chosen. This gave a wavelength of about 4 ft in a 20 mho/m solution. The pool dimensions in terms of water wavelengths then became 10 wavelengths long by 5 wavelengths wide.The 8-ft depth permitted simulation of an infinite medium for all practical purposes since the placement of a source at a 4-ft depth removed it a distance of one wavelength from all interfaces with an accompanying attenuation of approximately 55 db.

The sides of the pool were lined at the top with 1-ft wide strips of copper sheeting which were grounded in order to provide continuity between the conducting medium and ground. A satisfactory ground was attained by connecting the sheeting by means of braid to copper rods driven into the earth.

The pool was then filled with a 26 percent solution of sodium chloride giving a conductivity of approximately 20 mho/m for a temperature of 20° C.

A General Electric Navy Model TCU–2 transmitter fed the dipole and was operated at 296-kc continuous wave. The transmitter was placed in a double wall screen room located 35 ft from the dipole. To match the 100-ohm transmitter output to the approximately 1-ohm dipole, a transformer with an 11-to-1 turns ratio was used. The lead to the dipole consisted of 95-ohm twinax cable, RG 22/U, with shield grounded, encased in a copper tube which was buried in the earth and which lead directly into the conducting medium. The twinax cable in the conducting medium was connected to the transformer housed in a brass can which was supported at a height of 4 ft above the bottom of the pool. Electric dipoles consisting of insulated wire with silver-plated steel ball terminations were used. Dipole current was measured with a calibrated vacuum thermocouple housed in the brass can and connected in the secondary circuit. Single turn rectangular loops of insulated wire were also used as radiation sources. Magnetic field intensity was measured in air with a Radio Test Set AN/PRM–1.

5. Results

The dipoles and loops used as radiating sources are listed below. After a description of each source are listed the depth of immersion, the current fed to the source, the conductivity of the dissipative medium, and the component of magnetic field measured.

A. Horizontal 4-in. dipole; $h=2\frac{1}{8}$ in.; $I=3.77$ amp; $\sigma=18.2$ mho/m; measured $H\rho$.
B. Horizontal 4-in. dipole; $h=2\frac{1}{8}$ in.; $I=3.77$ amp; $\sigma=18.2$ mho/m; measured $H\phi$.
C. Horizontal 4-in. dipole; $h=6\frac{1}{8}$ in.; $I=3.77$ amp; $\sigma=18.2$ mho/m; measured $H\rho$.
D. Horizontal 4-ft dipole; $h=2\frac{1}{8}$ in.; $I=1.03$ amp; $\sigma=17.9$ mho/m; measured $H\rho$.
E. Horizontal 4-ft dipole; $h=38\frac{1}{8}$ in.; $I=1.26$ amp; $\sigma=16.1$ mho/m; measured $H\rho$.
F. Horizontal 3-ft dipole; $h=4$ in.; $Z=33\frac{1}{2}$ in.; $I=1.10$ amp; $\sigma=18.2$ mho/m; measured Hz.

G. Single turn 4-in. square loop with axis horizontal; $h=2$ in.; $h'=6$ in.; $I=3.77$ amp; $\sigma=18.2$ mho/m measured $H\rho$.
H. Single turn 4-ft by 3-ft loop with horizontal axis (4-ft top and bottom members); $h=2$ in.; $h'=38$ in.; $I=1.03$ amp; $\sigma=16.1$ mho/m measured $H\rho$.
I. U-Shape 4-in. square dipole; $h=6\frac{1}{8}$ in.; $I=3.77$ amp; $\sigma=18.2$ mho/m; measured $H\rho$.

All measurements with the above sources were made at a frequency of 296 kc.

Measured data are given in table 1 for the various sources. These data are compared in table 2, which shows the fall off and the experimental and calculated field values at 15 ft from each source. The calculated values were obtained by substitution into the appropriate eq (5), (7), or (11).

TABLE 1. *Measured magnetic field data using various radiating sources*

Distance from dipole in feet	Measured magnetic field intensity in amp/m$\times 10^5$								
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*
9	54.9	28.1	33.5	--------	3.95	--------	30.1	--------	33.4
11	30.8	16.3	20.4	--------	2.41	40.1	18.1	--------	20.4
13	21.6	9.37	13.4	70.2	1.55	22.1	11.5	83.6	12.8
15	14.7	6.82	8.70	47.5	1.07	13.4	7.63	55.0	8.69
17	10.0	5.08	6.15	34.8	0.682	9.03	5.55	40.1	6.02
19	6.89	3.75	4.48	25.4	.482	6.02	4.01	31.4	4.35
21	5.49	2.54	3.41	19.4	.335	4.01	3.01	22.7	3.41
23	4.35	2.41	2.54	14.7	.241	3.01	2.34	17.4	2.52
25	3.55	1.91	2.07	11.7	.174	2.17	1.87	13.4	2.01
27	2.51	1.51	1.61	8.90	--------	1.67	1.61	10.7	1.58
29	2.14	1.20	1.34	6.82	--------	1.27	1.20	8.56	1.32

TABLE 2. *Comparison of magnetic field data from various radiating sources*

Source	Measured fall off	Measured field at 15 ft *amp/m* $\times 10^5$	Calculated field at 15 ft *amp/m* $\times 10^5$
A	2.9	14.7	15.3
B	3.1	6.82	7.65
C	3.0	8.70	9.59
D	3.0	47.5	50.7
E	3.3	1.07	1.03
F	3.9	13.4	10.2
G	3.0	7.63	8.20
H	3.0	55.0	55.9
I	3.0	8.69	9.37

The data from sources *A* and *B* are plotted in figure 2. This plot shows the inverse cube fall off for $H\rho_2$ and $H\phi_2$ and the 6-db factor between them.

An examination of the results given in table 2 shows good agreement with theory both with respect to fall off and absolute value. All experimental field values except those from source *F* were within 1 db of the calculated values. The field value from source *F* was within 2.4 db of the calculated value.

The nulls of figure 3 were found to occur at 35°32′ for the 20-ft traverse and 37°11′ for the 55-ft traverse, thus verifying eq (4).

From the experimental and calculated results presented, it appears that the field is determined solely by the current in the horizontal radiating wires of the dipole or loop. Since this conclusion is not particularly obvious, the following experiment was performed to more clearly determine whether or not this was actually the case.

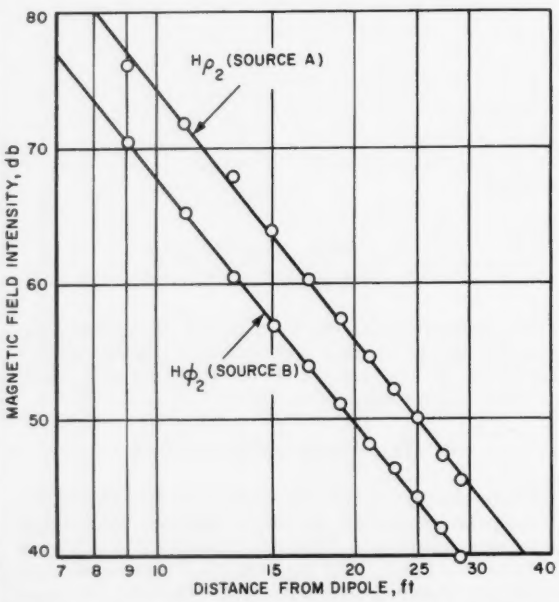


FIGURE 2. *Fall off of horizontal magnetic field components.*
Reference level: 9.65×10^{-8} amp/m

Sources *A*, *C*, and *I* were placed below the surface of the conducting medium as shown in figure 4, and the field radiated from each was measured. The

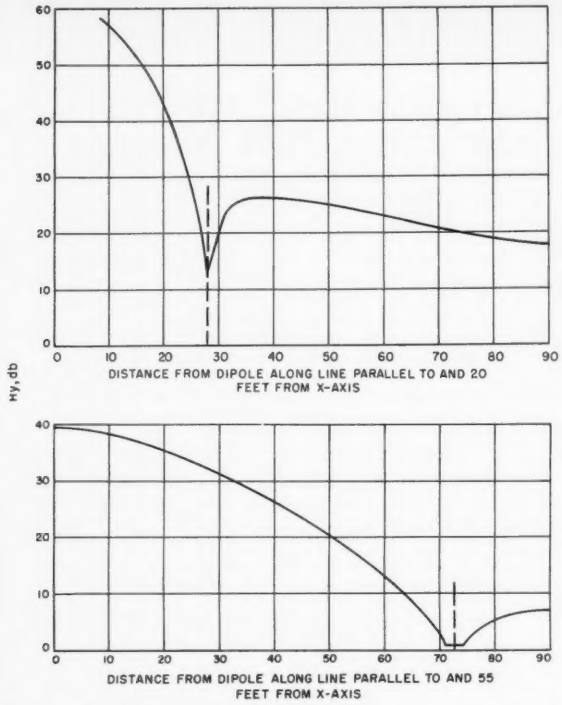

Hy, db
DISTANCE FROM DIPOLE ALONG LINE PARALLEL TO AND 20 FEET FROM X-AXIS
DISTANCE FROM DIPOLE ALONG LINE PARALLEL TO AND 55 FEET FROM X-AXIS

FIGURE 3. *Magnetic field plot of* Hy *from traverse parallel to dipole axis.*

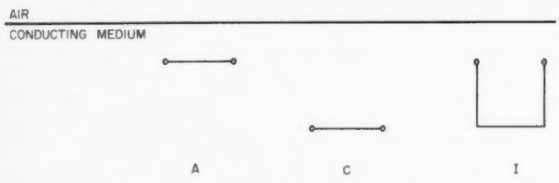


FIGURE 4. *Positions of radiating sources for determining localization of current moment.*

results given in tables 1 and 2 show that source I was equivalent to source C and not to source A. Thus, the currents in the horizontal wires of the dipoles determine the value of the radiated field.

6. Conclusions

On the basis of the experimental results described, the following conclusions may be drawn:

(a) The magnetic field propagation eq (1), (2), and (3), for the range therein stated, are valid.

(b) For calculating the magnetic field of a dipole using these equations, the current moment may be considered as due only to the currents in the horizontal wires forming the dipole.

(c) Calculations for a rectangular loop with horizontal axis may be made on the basis of the loop consisting of two horizontal dipoles carrying current in opposite directions and properly phased according to depth in the conducting medium.

7. References

[1] C. T. Tai, Hertzian dipole immersed in a dissipative medium, Cruft Lab. Rept. No. 21, Harvard Univ. (1947).

[2] R. K. Moore, The theory of radio communication between submerged submarines, Ph. D. Thesis, Cornell Univ. (1951).

[3] W. Von Aulock, Low frequency electromagnetic dipole fields in a semi-infinite conductor, Navy Dept., Bureau of Ships, Minesweeping Sec., Tech. Rept., No. 104 (1952).

[4] J. R. Wait, Magnetic dipole antenna in a conducting medium, Proc. I.R.E., **40,** 1244–45 (1952).

[5] J. R. Wait and L. L. Campbell, The fields of an oscillating magnetic dipole immersed in a semi-infinite conducting medium, J. Geophys. Research, **58,** 167–168 (1953).

[6] R. H. Lien and J. R. Wait, Radiation from a horizontal dipole in a semi-infinite dissipative medium, J. Appl. Phys. **24,** 1–5, 958–959 (1953).

[7] A. Baños, Jr., and J. P. Wesley, The horizontal electric dipole in a conducting half-space, Pt. I, Univ. Calif., Marine Phys. Lab., Scripps Inst. Oceanography, SIO Ref. 53–33 (Sept. 1953); Pt. II, SIO Ref. 54–31 (Aug. 1954).

BOULDER, COLO. (Paper 64D1–33)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

A Very-Low-Frequency Antenna for Investigating the Ionosphere With Horizontally Polarized Radio Waves[1]

R. S. Macmillan,[2] W. V. T. Rusch, and R. M. Golden[3]

(August 4, 1959)

The advantages of a horizontal half-wave resonant antenna for very-low-frequency propagation experiments lie in its relatively simple and inexpensive construction and in its radiation pattern which is maximum in the vertical direction. The radiation fields of this type of antenna located at the surface of a conducting earth consist of: 1, A horizontally polarized space-wave field radiated in the perpendicular bisector plane of the antenna; and 2, a vertically polarized groundwave field radiated along the axis of the antenna. This vertically polarized field is zero at right angles to the antenna. These fields have been experimentally verified.

The use of a 50-kilocycle horizontal half-wave antenna for vertical-incidence ionospheric sounding experiments is described. The radiation pattern of this antenna is well suited for ionospheric soundings since a receiver located in the groundwave null receives only the reflected skywave signal.

Ground-resistivity measurements made at a number of locations in Central and Southern California were correlated with the geology of the terrain. This correlation showed that the ground resistivity is highest (a condition necessary for optimum antenna efficiency) in areas where the underlying rock formations are relatively unfractured. The amount of annual rainfall and other climatic conditions have little or no effect on the resistivity.

Finally, a unique antenna system is presented which employs resonant loading circuits to convert a section of an existing power line into a horizontal half-wave very-low-frequency transmitting antenna.

1. Introduction

A half-wave, horizontal, linear antenna mounted above the ground produces a horizontally polarized radiation field. This type of antenna is relatively inexpensive to construct, requires no tuning elements or ground system, and is well suited for many vlf propagation experiments. It is the purpose of this paper to analyze the radiation properties of this antenna, and to compare these theoretical results with experimental observations. The methods used can be generalized to certain types of more complicated antenna systems. The effects of the underlying geology upon the radiation properties of the antenna have also been considered, and the conclusions reached should be of considerable value in the selection of vlf antenna sites. Finally, a unique system is presented whereby a commercial power line is used as a vlf transmitting antenna.

2. Antenna Radiation Fields

The problem of an oscillating dipole located above a plane earth of finite conductivity is of great interest to radio engineers. The classical investigation of this problem was published by Arnold Sommerfeld [1],[4] who expanded the Hertz vectors from both the primary excitation and the induced secondary effects of the ground in a continuous spectrum of eigenfunctions, and then matched tangential fields at the boundary. In order to match these fields for a dipole parallel to the earth, Sommerfeld has demonstrated the necessity for two components of the Hertz vector, π_x and π_z. Using the coordinates of figure 1:

$$\pi_x=-\frac{e^{ikR_1}}{R_1}+\frac{e^{ikR_2}}{R_2}-\int_0^\infty \frac{2J_0(\lambda\rho)e^{-(h+z)\sqrt{\lambda^2-k^2}}\lambda d\lambda}{\sqrt{\lambda^2-k^2}+\sqrt{\lambda^2-k_E^2}}, \quad (1)$$

$$\pi_z=-\cos\phi\int_0^\infty \frac{2(1-u^2)J_0'(\lambda\rho)e^{-(h+z)\sqrt{\lambda^2-k^2}}\lambda^2 d\lambda}{(\sqrt{\lambda^2-k^2}+\sqrt{\lambda^2-k_E^2})(\sqrt{\lambda^2-k^2}+u^2\sqrt{\lambda^2-k_E^2})} \quad (2)$$

where $u^2=k^2/k_E^2$, $k^2=\omega^2\mu_0\epsilon_0$, $k_E^2=\omega^2\mu\epsilon+i\omega\mu\sigma$, and the path of the integration is taken along the real axis of the complex λ-plane. Similar results were obtained by H. Weyl [2] who recognized that the problem could be treated as the reflection and refraction of bundles of plane waves for different angles of incidence at the earth's surface. Unfortunately, (1) and (2) are not readily adapted to numerical computation, and in the years following Sommerfeld's original work, many researchers have attempted to reduce them to more manageable form. A survey of the entire problem is made by J. A. Stratton [3].

[1] This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under contract No. AF 18(600)-1552, with the exception of the research described in the section on experimental data, which was supported by the Office of Naval Research under contract Nonr.-220(07). The main results of the paper were presented at the Symposium on VLF Radio Waves held in Boulder, Colo., January 1957.

[2] University of Southern California, Los Angeles, California.
[3] California Institute of Technology, Pasadena, California.

[4] Figures in brackets indicate the literature references at the end of this paper

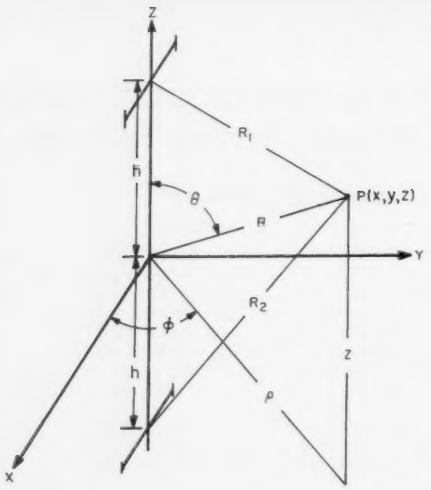


FIGURE 1. *Geometry.*

K. A. Norton [4] has computed the space and ground waves in the radiation field (i.e., $R>>\lambda$) of a horizontal electric dipole located on the surface (i.e., $h\approx 0$) of a lossy plane earth. From Norton's results, the electric space-wave radiation field in the plane at right angles to the axis of the dipole is:

$$E_x=-ik\left[\frac{I\eta dx}{4\pi}\right]\left[\frac{2\cos\theta}{\sqrt{1/u^2-\sin^2\theta}+\cos\theta}\right]\left[\frac{e^{i(kR-\omega t)}}{R}\right], \tag{3}$$

where I is the dipole current, η is 120 π, the intrinsic impedance of free space, and dx is the length of the dipole.

In order to compute the corresponding field from a horizontal half-wave antenna, the contributions from each infinitesimal element are integrated along a current distribution $I_o \cos kx$. (The justification for assuming this current distribution is presented below.) Then, for the half-wave horizontal antenna located on the surface of a lossy plane earth, the electric radiation field in the plane $\phi=\pi/2$ is:

$$E_x=\frac{-i\eta I_o}{2\pi}\left[\frac{2\cos\theta}{\sqrt{1/u^2-\sin^2\theta}+\cos\theta}\right]\left[\frac{e^{i(kR-\omega t)}}{R}\right]. \tag{4}$$

This horizontally polarized field becomes zero at the surface of the earth.[5] Its radiation pattern (fig. 2) has been verified experimentally (see below) and degenerates into the correct expansion for the equivalent field of a free-space half-wave antenna when $u=1$.

[5] The horizontally polarized groundwave actually varies as $1/R^2$ but is negligible at vlf [5].

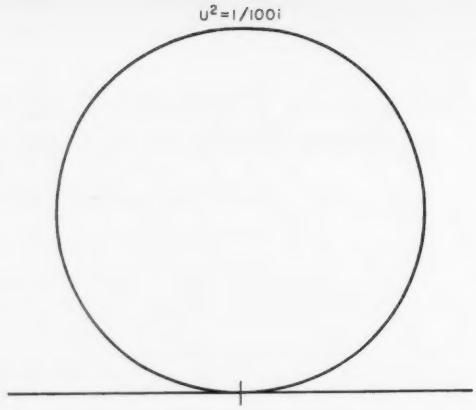


FIGURE 2. *Radiation pattern in plane* $\phi=\pi/2$.

From Norton's results, the electric space-wave radiation field in the vertical plane ($\phi=0$) of a horizontal electric dipole is:

$$E_\theta=-ik\left[\frac{I\eta dx}{4\pi}\right]\left[\frac{2u\cos\theta\sqrt{1-u^2\sin^2\theta}}{\cos\theta+u\sqrt{1-u^2\sin^2\theta}}\right]\left[\frac{e^{i(kR-\omega t)}}{R}\right]. \tag{5}$$

For a half-wave antenna it is again only necessary to integrate (5) along a half-wave cosinusoidal current distribution to find the field at any angle θ. However, in this case the integration must account for the phase differences in the contributions from different sections of the antenna. R is therefore replaced by $R-x\sin\theta$ in the exponent. Then, for the horizontal half-wave antenna located on the surface of a lossy earth, the electric radiation field in the plane $\phi=0$ is:

$$E_\theta=\frac{-i\eta I_0}{\pi}\left[\frac{u\sqrt{1-u^2\sin^2\theta}}{\cos\theta+u\sqrt{1-u^2\sin^2\theta}}\right]\times\left[\frac{\cos\left(\frac{\pi}{2}\sin\theta\right)}{\cos\theta}\right]\left[\frac{e^{i(kR-\omega t)}}{R}\right]. \tag{6}$$

The radiation pattern of this field in the vertical plane of the antenna is plotted in figure 3. Equations (6) and (4) are identical when $\theta=0$.

At the surface of the earth the field of a horizontal electric dipole consists entirely of the groundwave since the space wave becomes zero when $\theta=\pi/2$. Conversely, the groundwave rapidly becomes negligible with increasing height. The dominant component of the groundwave of the horizontal electric dipole is verticallly polarized:

$$E_z=ik\left[\frac{I\eta dx}{4\pi}\right][\cos\phi u\sqrt{1-u^2}\,2F]\left[\frac{e^{i(kR-\omega t)}}{R}\right] \tag{7}$$

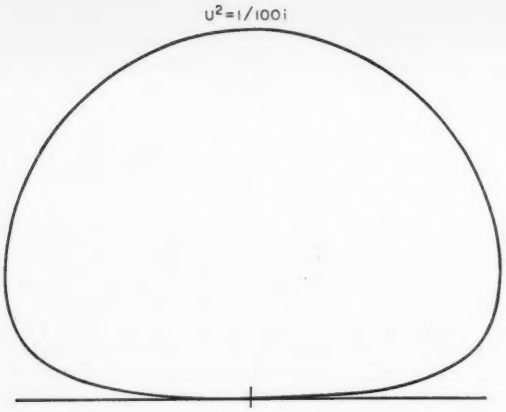


FIGURE 3. *Radiation pattern in plane* $\phi=0$.

where

$$F=1+i\sqrt{\pi w}\, e^{-w}\, \mathrm{cerf}\,(-i\sqrt{w}) \tag{8}$$

and

$$w=\frac{ikRu^2(1-u^2)}{2}\ (|w| \text{ is the numerical distance}). \tag{9}$$

There is a radial component of the groundwave field, smaller than the vertically polarized component by a factor of $u\sqrt{1-u^2}$ (the "wave-tilt"). There is also a ϕ-component, smaller than the vertical component by a factor of u^3 at large distances, which can be neglected for practical values of earth conductivity.

Integration of (7) along the antenna yields the vertically polarized groundwave field in the plane $\theta=\pi/2$. In evaluating the integral, R is replaced by $R-x\cos\phi$ in the exponent. The groundwave field of a horizontal half-wave antenna located at the surface of a lossy plane earth is then:

$$E_z=\frac{i\eta I_0}{\pi}\left[u\sqrt{1-u^2}\right]\left[\frac{\cos\phi\cos\left(\frac{\pi}{2}\cos\phi\right)}{\sin^2\phi}\right]$$

$$\times\left[F\right]\left[\frac{e^{i(kR-\omega t)}}{R}\right]. \tag{10}$$

This vertically polarized groundwave field in the horizontal plane of the antenna is maximum along the axis of the antenna and zero at right angles to the antenna. The radiation pattern of this vertical field is plotted in figure 4. Horizontal components of the groundwave field can be neglected for small values of the wave-tilt. Equation (10) was also derived by Wait [5].

Examination of (7) yields information on the properties of the groundwave. For large values of R, F asymptotically approaches $-\frac{1}{2}\,w$, and hence the groundwave ultimately varies inversely as the square of the distance. In the vicinity of the antenna, the value of F is unity, and the groundwave varies inversely as the first power of the distance. A complete discussion of the radial attenuation properties of the groundwave is presented in [4] and [6].

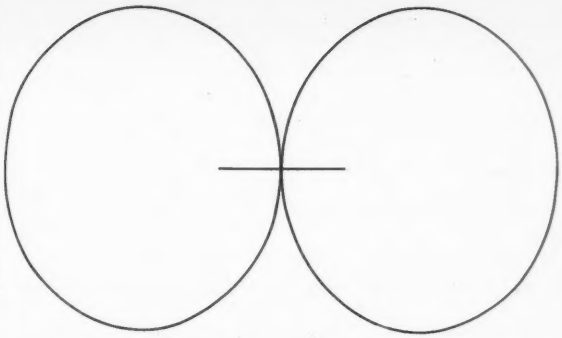


FIGURE 4. *Surface-wave radiation pattern in plane* Z=0.

The fields in a plane perpendicular to the center of the antenna ($\phi=\pi/2$ plane) can also be obtained by considering the superposition of the directly radiated fields and the fields reflected from the plane of the earth.

From the expression for a half-wave antenna in free space, it is seen that the entire electric field is perpendicular to the plane $\phi=\pi/2$. Hence it is possible to apply the laws of reflection and refraction to this field. This method can be applied to many more complicated field problems, which would otherwise be difficult to solve.

The geometrical configuration for the path of the direct and reflected fields at distances great compared to the height of the antenna is shown in figure 5(a). By properly matching the fields at the surface of the earth, and neglecting the attenuation due to the small difference in the path lengths, one obtains an expression for the total electric field of the form

$$E_{\text{total}}=(1+\rho e^{-i\delta})E_{\text{direct}}, \tag{11}$$

where

E_{direct} is the free space field in the plane $\phi=\pi/2$ of a half-wave antenna and equals

$$\frac{\omega\mu_0 I_0}{2\pi kR}.$$

$$\rho=\frac{u\cos\theta-\sqrt{1-u^2\sin^2\theta}}{u\cos\theta+\sqrt{1-u^2\sin^2\theta}},\ \text{reflection coefficient.}$$

$$\delta=\frac{4\pi h}{\lambda_0}\cos\theta,\ \text{phase delay of reflected wave.}$$

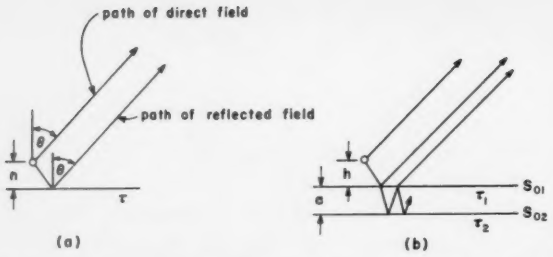


FIGURE 5. *Geometrical configuration for the reflection method.*

For cases in which the height of the antenna is small compared to a free space wavelength, the ratio of the total electric field in the plane $\phi=\pi/2$ to the corresponding field in free space is given approximately by:

$$\frac{E_{\text{total}}}{E_{\text{direct}}}=1+\rho=\frac{2\cos\theta}{\cos\theta+\sqrt{1/u^2-\sin^2\theta}}. \qquad (12)$$

Equation (12) is the same as the previously derived eq (4).

This derivation has assumed a perfectly homogeneous earth. In many locations, however, rock of high resistivity is covered by a mantle of low resistivity soil. Hence this case will now be considered.

A homogeneous earth is assumed to be covered with a layer of thickness "a" which has a much lower resistivity than the material that it covers, as shown in figure 5(b).

The reflection method is still useful for this case since the infinite series obtained due to multiple reflections and refractions at the two surfaces S_{01} and S_{12} can be readily summed. One thus obtains for the ratio of the total field to the direct field:

$$\frac{E_{\text{total}}}{E_{\text{direct}}}=1+e^{-i\delta}\left[\frac{\rho_{01}+\rho_{12}e^{-i\delta_a}}{1+\rho_{01}\rho_{12}e^{-i\delta_a}}\right], \qquad (13)$$

where

$$\rho_{01}=\frac{u_1\cos\theta-\sqrt{1-u_1^2\sin^2\theta}}{u_1\cos\theta+\sqrt{1-u_1^2\sin^2\theta}}$$

reflection coefficient for waves reflected from S_{01}.

$$\rho_{12}=\frac{u_2\sqrt{1-u_1^2\sin^2\theta}-u_1\sqrt{1-u_2^2\sin^2\theta}}{u_2\sqrt{1-u_1^2\sin^2\theta}+u_1\sqrt{1-u_2^2\sin^2\theta}}$$

reflection coefficient for waves reflected at S_{12}.

$$\delta_a=(4\pi a/\lambda_c)\,\frac{\sqrt{1-u_1^2\sin^2\theta}}{u_1}$$

phase delay in top layer.

Expressing u_1 in terms of the skin depth, d_1, in the top layer gives $u_1=kd_1/2$. For a resistivity of 100-ohm-m, and for frequencies between 5 and 20 kc, the skin depth varies from 70 to 35 m. Typical values for the thickness of the top layer are several meters. If the thickness is small compared to a skin depth, the ratio of the fields is very nearly

$$\frac{E_{\text{total}}}{E_{\text{direct}}}=1+\frac{\rho_{01}+\rho_{12}}{1+\rho_{01}\rho_{12}}=\frac{2\cos\theta}{\cos\theta+\sqrt{1/u_2^2-\sin^2\theta}}, \qquad (14)$$

which is the same as expression (12) derived for the case of the homogeneous earth. Thus, the influence of a good conducting layer can be neglected if its thickness is very much less than a skin depth.

The magnitude of the ratio of the total field directly above the antenna to the free-space field for a relative capacitivity of 6 is plotted in figure 6 as a function of the product of frequency and resistivity.

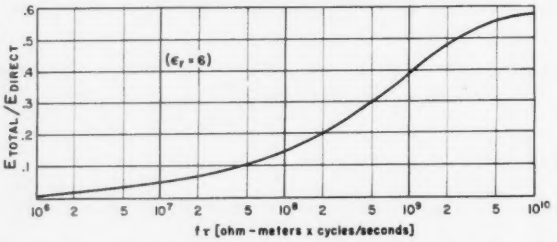


FIGURE 6. *Magnitude of ratio of total field directly above the antenna to free-space field.*

For the case of a stratified earth, the reflection method becomes cumbersome. This case is most readily solved by the wave method in which the field is resolved into traveling waves. By assuming solutions in each layer and matching the fields at the interfaces, one can obtain solutions for the fields directly above a horizontal antenna of arbitrary length. Wait has applied this method to a long wire filament and has obtained an expression for the electric field on the surface of the stratified medium [7].

2.1. Experimental Results

The properties of a horizontal half-wave antenna were investigated experimentally with three antennas located over ground with varying electrical properties.

The first antenna, half-wave resonant at 197 Mc, was located over dry rocky soil. Field strength measurements were made in the perpendicular bisector plane of the antenna. Figure 7 shows the theoretical and measured radiation patterns in this plane. By noting the ratio of maximum to minimum field strength directly over the antenna as the height of the antenna was varied, it was possible to calculate the relative capacitivity of the underlying ground. From (11):

$$\frac{E_{\max}}{E_{\min}}=\frac{1-\rho}{1+\rho}=\frac{1}{u}=\sqrt{\epsilon_r},\ (\rho<0). \qquad (15)$$

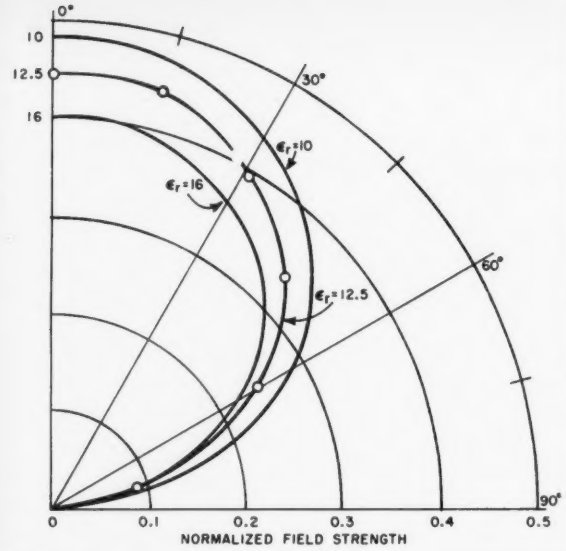


FIGURE 7. *Field pattern of* λ/2 *antenna over a plane earth* h=0 *(calculated results for* ϵ_r=*16, 10 and measured values for* ϵ_r=*12.5).*

Using a value of 12.5 for the relative capacitivity it is seen that the radiation pattern of the antenna is in good agreement with the theoretical result given by (4).

The input resistance of the antenna was also measured as a function of antenna height. From figure 8 it is seen that the resistance fluctuated about 70 ohms when the height was greater than λ/4, but increased rapidly when the height approached zero. At a height of 1.5 cm the input resistance was 80 ohms. The rise in resistance may be attributed to increased ground losses.

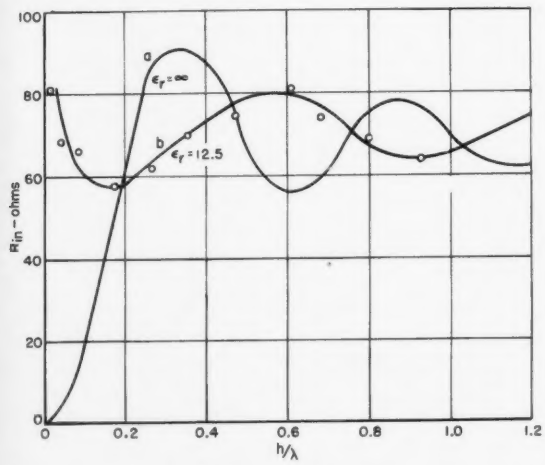


FIGURE 8. *Input resistance of a half-wave antenna above a plane earth.*
(a) Calculated for ϵ_r=∞
(b) Experimental for ϵ_r=12.5.

A second antenna approximately 2,000 m in length was constructed above ground of relative capacitivity ϵ_r=6 or 7 and resistivity 137 ohm-m. The average height of the antenna was 1 m and it was resonant at 60 kc. The input resistance of the antenna (including 24 ohms copper loss) was 127 ohms. Changing the average height of the antenna to 3.4 m reduced the input resistance to 79 ohms. This change in resistance may be attributed almost entirely to the reduction in ground losses.

The third antenna was erected at approximately the same site as the second antenna and consisted of a wire 8,000 m long, supported about 3.4 m above the ground. By locating switches along the wire, it was possible to tune the antenna to different frequencies (simply by opening the appropriate pair of switches). The antenna was thus half-wave resonant in the frequency range between 18 kc and 250 kc. The vertically polarized field at the surface of the ground was in excellent agreement with the theoretical pattern of figure 4. The wave tilt was found to be very small as expected for this type of terrain. No vertically polarized field could be detected at the surface of the ground in the equatorial plane of the antenna.

The input resistance to the antenna at 18 kc was 82 ohms, 20 ohms of which was the estimated copper loss. The quality factor, Q, of the antenna was measured to be 12 at this frequency.

Several continuous-wave, vertical-incidence, ionospheric-sounding experiments using this antenna were conducted by Bergman et al. [8]. The antenna was driven by a 800-w transmitter which was tunable over a frequency range from 14 to 70 kc. The receiving site was located at a distance of 64 km from the transmitter in the perpendicular bisector plane of the transmitting antenna. Since no ground wave existed in this plane, it was possible to use two crossed dipoles to directly detect the relative amplitude and phase of the downcoming skywave, without subtracting a ground-wave component.

The dipoles were each 30 m in length and were oriented parallel and perpendicular to the vertical plane of the transmitting antenna. This orientation made it possible to record directly the skywave components polarized parallel to and normal to the plane of incidence.

By transmitting a continuous signal, data on the variations in height of the ionosphere were recorded. This was accomplished by comparing the phase of the received signal with a reference phase signal transmitted over a direct line-of-sight vhf communication channel.

The geometry of the sites is shown in figure 9. The particular spacing and orientation was governed by the geography of the area, and the requirement of a line of sight between stations. A block diagram of the complete system is shown in figure 10.

The experimental antennas thus made it possible to verify the theoretical radiation fields of a half-wave antenna from eqs (4) and (10). The rapid increase in antenna resistance as the antenna height approached zero is attributed to the increase in ground losses. This result was predicted by Sommerfeld and

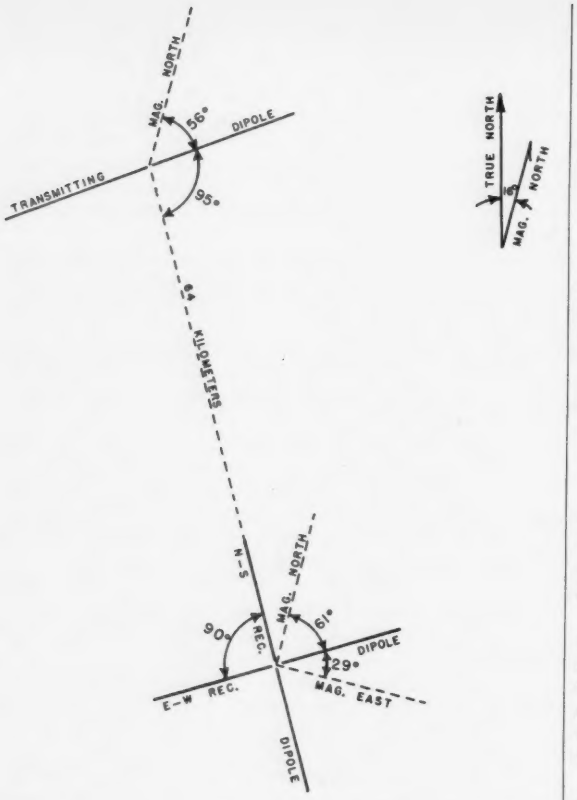


FIGURE 9. *Geometry of the transmitting and receiving antennas.*

Renner [9] who determined the radiation and earth absorption of a dipole antenna over a finite conducting earth. Their results are basically those presented in figure 8. This rapid increase is due to the extremely high fields in the immediate vicinity of the antenna wire. Thus, in order to prevent excessive losses, a horizontal vlf antenna should be placed at least several meters above the earth's surface.

3. Form of the Current Distribution

Coleman [10] has shown that the propagation constant of waves traveling along a wire located at the interface between two media is strongly affected by the intrinsic propagation constants of the two media. In the case of actual ground constants encountered at very low frequencies, the propagation constant for waves traveling along a wire located at an earth-air interface is

$$\Gamma \cong \frac{1}{2}(1+i)\sqrt{\sigma_2 \mu \omega}. \tag{16}$$

However the change in propagation constant from the free-space value to the interface value is expected to be quite sharp, the transition taking place in a few diameters of the wire [5].

The horizontal, half-wave antenna discussed in this paper was mounted on poles, several meters above the earth. The fields of eqs (4), (6), and (10) were calculated assuming the antenna to lie on the earth (i.e., $h=0$). This approximation is justified when it is considered that at very low frequencies the wavelength is at least 10 km, so that $\lambda >> h$.

On the other hand, the criterion that the antenna wire be more than a few wire diameters above the earth is also satisfied, i.e., $d << h$. Since the prop-

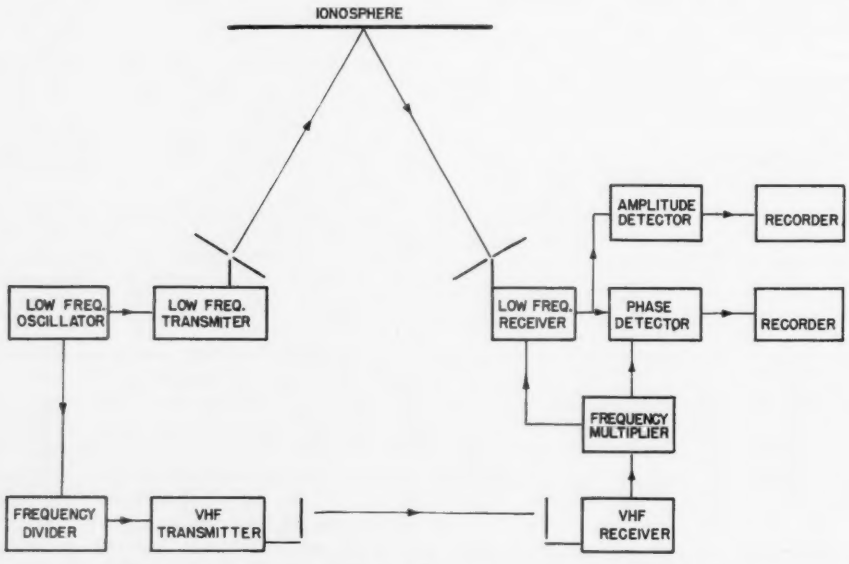


FIGURE 10. *Block diagram of the transmitting and receiving system.*

agation constant should then be close to that of free space, the approximation has been made that the current distribution on the horizontal, half-wave antenna is $I_{rms}(x)=I_0 \cos kx, -\frac{\lambda}{4} \leq x \leq \frac{\lambda}{4}$. To determine the validity of this approximation, an experiment was performed using the antenna resonant at 60 kc. The antenna was approximately 2,000 m long and had an average height of 1 m. The wire diameter was approximately ½ cm. The antenna was driven at its center with a power oscillator, and the antenna current distribution was determined by inserting a thermal ammeter at a number of points along the antenna wire. The accuracy of the current distribution which was obtained was limited by the linearity of the ammeter and any small fluctuations of the oscillator power level which may have occurred during the course of the measurements. Within the estimated accuracy of 5 percent for the experiment, no departure from a cosinusoidal current distribution could be detected.

A second experiment was performed with a 10 Mc antenna mounted 1 m above the ground with virtually the same results as for the 60 kc antenna.

Furthermore, for the high values of ground resistivity which are of interest, and with the antenna located several meters above the earth's surface, any departure from the cosinusoidal current distribution should be pronounced only towards the ends of the antenna where the magnitude of the current is small. The authors believe that the departure from the assumed cosinusoidal current distribution would have only a very minor if not negligible effect on the values of the radiation fields of eqs (4), (6), and (10).

4. Ground Resistivity

For antennas of the type being considered, it has been shown that the product of the earth's resistivity and the excitation frequency must be high to obtain a reasonable value for the radiated fields (see fig. 6). In order to find a site with extremely high ground resistivity, measurements were made at a number of locations in Central and Southern California (see fig. 11). The resistivity measurements were made by the four-electrode method [11]. The four electrodes are placed in the ground at specified distances

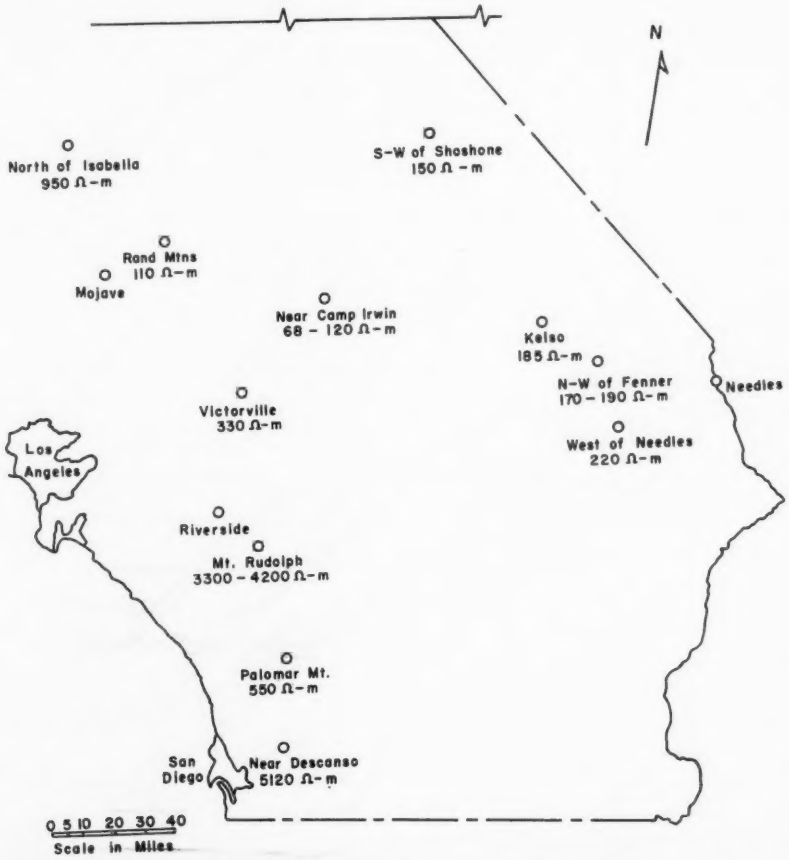


FIGURE 11. *Southern California ground resistivity measurements.*

apart, and a current is passed between the outer electrodes while the potential difference between the inner two is measured. From this measurement, the ground resistivity can be determined.

The resistivity measurements were made in three general areas of California—the Mojave Desert region, a triangular section (indicated by the numeral II in fig. 12) of the Peninsular Range province, and a section of the Sierra Nevada Mountains east of Fresno.

The Mojave Desert is a broad interior region of isolated mountain ranges separated by expanses of desert plains. It has an enclosed drainage system with playas, except for the bordering Colorado River province on the east. There are two important fault trends: the NW-SE trend is the more prominent, while the E-W trend is secondary. The province is wedged in a sharp angle between the Garlock fault and the San Andreas fault (see fig. 12).

The Peninsular Range province consists of a series of ranges separated by longitudinal valleys, trending NW-SE. The trend of the topography is like that of the coast ranges, but the geology is more like that of the Sierra Nevada. The dominating rocks are granitic, intruded into older metamorphic series. The coastal area receives 11 to 18 in. of rainfall per year while some of the higher mountains receive as much as 50 in./yr.

The Mojave Desert resistivity measurements were made in areas which: (1) Received very little rainfall; (2) had a very-low water table (as low as 2,300 ft below the surface); and (3) were underlaid by granitic rocks. However, the resistivity values were much lower than expected. These values ranged from 68 to 330 ohm-m (see fig. 11).

The low values of resistivity are attributed to the extensive faulting which occurs in the province (see fig. 12). The rocks are badly fractured, and these fractures contain water which is held by capillary forces. Thus the resistivity is low, even though the area receives little rainfall—as low as 0.5 to 2 in./yr in some sections.

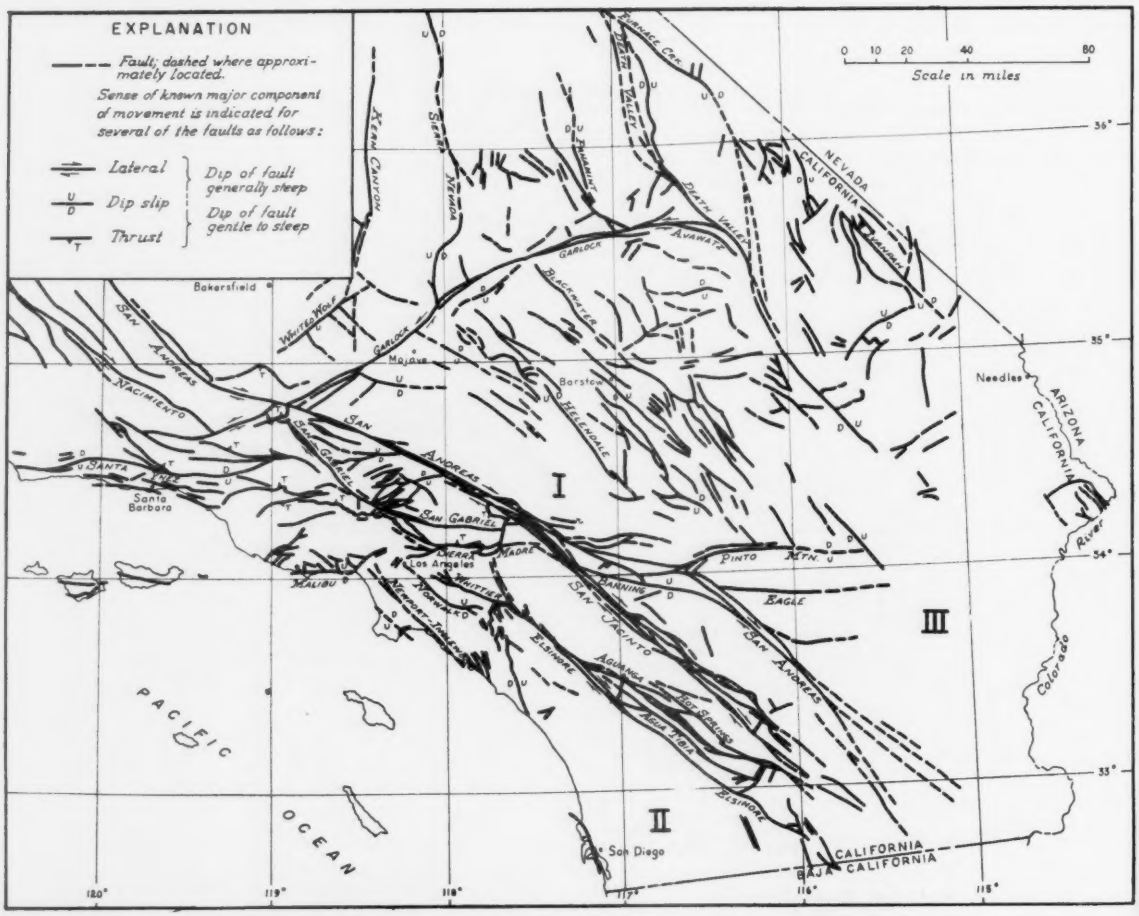


FIGURE 12. *Map showing major faults in a large part of Southern California.*
Based mainly on mapping and compilation by C. R. Allen, T. L. Bailey, T. W. Dibblee, Jr., D. F. Hewett, R. H. Jahns, and L. F. Noble.

The triangular section indicated by the numeral II in figure 12 is a large block or batholith which is free of faulting. Large areas of relatively unfractured granite were found in both this triangular section and the Sierra Nevada Mountains. The resistivity values for these areas were found to range from 2,100 to 5,120 ohm-m. It is concluded that for high values of resistivity, the rock formations must be relatively unfractured, while the amount of annual rainfall and other climatic conditions are of little significance.

It might be inferred from figure 12 that there is little faulting in the areas indicated by the numerals I and III. However, this is not the case. The faults are not shown in these areas, because deep aluvium deposits prevent their accurate location.

5. Power-Line Antenna

In 1955 a program of research was initiated by the California Institute of Technology under the sponsorship of the Office of Scientific Research of the Air Research and Development Command. The purpose of this research was to investigate various aspects of vlf propagation.

Initially it was proposed to construct a horizontal half-wave antenna. It was subsequently proposed, however, that a commercial power line might be used as the transmitting antenna. This proposal led to the analysis, design, and construction of a unique vlf antenna system. This system employed a single-phase, medium voltage, commercial power line. The antenna section was isolated from the rest of the distribution system by parallel resonant circuits, or line-traps (see fig. 13). Two of these traps were located at each end of the half-wavelength antenna section, and two at the center of the antenna. The end traps served principally to prevent the radiofrequency energy from coupling past the antenna section into the remainder of the power distribution system. The center traps isolated the antenna section into two halves so that it could be fed as a balanced, center-fed antenna. The radiofrequency traps were designed to present a negligible impedance to the flow of 60 cps power and a maximum impedance to the radiofrequency power. In this manner it was possible to draw 50 kw of 60 cps power from the center of the antenna section to drive the transmitter while simultaneously radiating vlf energy from the same section of line.

The radiation patterns of the power-line antenna

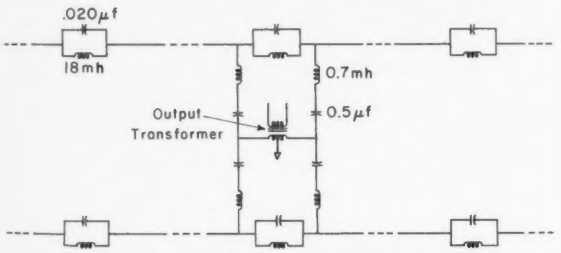


FIGURE 13. *Schematic diagram of the power-line antenna system.*

are virtually identical to those of figures 2, 3, and 4 [12].

The power-line antenna was used primarily for ionospheric sounding experiments at 8.4 kc. Measurements of the amplitude variations of the reflected sky wave signal revealed information about the lower regions of the ionosphere. Transcontinental detection of the 8.4-kc signal from California to Cambridge, Mass., indicated that the antenna system is also capable of long-range propagation.

6. Conclusions

The radiation field of a half-wave, linear antenna located above a plane earth of finite conductivity consists of: (1) A horizontally polarized field in the midplane perpendicular to the antenna, which is zero in the plane of the earth; and (2) a vertically polarized groundwave which is maximum in the direction of the axis of the antenna and zero at right angles to the antenna. This radiation field, given in eq (4), (6), and (10) has been confirmed by experimental data. The nature of its radiation field makes this simply constructed antenna extremely useful in ionospheric propagation experiments. The radiation is maximum in the vertical direction, and in the midplane of the antenna there is no direct groundwave radiation to interfere with the reflected skywave information.

To obtain maximum radiation fields, it is necessary to locate the antenna over ground of high resistivity. This type of ground is primarily located in geological areas free of faulting.

7. References

[1] A. Sommerfeld, The propagation of waves in wireless telegraphy, Ann. Physik. **28,** 665 (March 1909).

[2] H. Weyl, The propagation of electromagnetic waves over a plane conductor, Ann. Physik. **60,** 481 (Nov. 1919).

[3] J. A. Stratton, Electromagnetic theory, 573 (McGraw-Hill Book Co. Inc., New York, N.Y., 1941).

[4] K. A. Norton, The physical reality of space and surface waves in the radiation field of radio antennas, Proc. IRE **25,** 1192 (Sept. 1937).

[5] J. R. Wait, Radiation from a ground antenna, Can. J. Technol. **32,** 1 (Jan. 1954); also, Low-frequency radiation from a horizontal over a spherical earth, Can. J. Phys. **34,** 586 (1956).

[6] J. R. Wait, Excitation of surface waves on conducting stratified dielectric-clad and corrugated surfaces, J. Research NBS **59,** 365 (Dec. 1957) RP2807.

[7] J. R. Wait, The fields of a line source of current over a stratified conductor, Appl. Sci. Research **[B] 3,** The Hague, Martinus Nijhoff (1953).

[8] C. W. Bergman, R. S. Macmillan, W. H. Pickering, A new technique for investigating the ionosphere at low and very low frequencies, (Rocket Exploration of the Upper Atmosphere, Interscience Publishers, Inc., New York, N.Y., 1954).

[9] A. Sommerfeld, F. Renner, Strahlungsenergie und erde absorption bei dipole antennen, Ann. Physik. **41,** 1 (Jan. 1942).

[10] B. L. Coleman, Propagation of electromagnetic disturbances along a thin wire in a horizontally stratified medium, Phil. Mag. **41,** 276 (1950).

[11] W. R. Smythe, Static and dynamic electricity, 241 (McGraw-Hill Book Co., Inc., New York, N.Y., 1950).

[12] W. V. T. Rusch, A new transmitting antenna system for very low radio frequencies, Ph.D. Thesis, Calif. Inst. Technol. (1959).

BOULDER, COLO. (Paper 64D1–34)

Journal of Research of the Naticnal Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Effects of High-Altitude Nuclear Explosions on Radio Noise

C. A. Samson

(July 9, 1959)

High-altitude nuclear explosions over Johnston Island in August 1958 appear to have had a rather pronounced effect on the radio noise recorded at Kekaha, Hawaii. Graphs are presented showing the hour-to-hour variation of the noise during August at eight frequencies from 13 kc to 20 Mc. All frequencies seem to have been affected, and the drop in received noise power amounted to as much as 32 db in the hour following the first explosion. The period of time over which abnormal noise conditions were observed suggests that high-altitude nuclear explosions may have a rather persistent effect on radio communications at certain frequencies.

1. Introduction

Atmospheric radio noise originates primarily in thunderstorms, and the amount of noise received at a given location is influenced not only by storms in the immediate vicinity, but also by storms a few hundred to several thousand miles distant. Since the noise from these distant storms is propagated by the ionosphere, a record of the radio noise will give an indication of changes in the ionosphere, particularly if there is little or no obscuring effect from storms close enough for groundwave propagation.

Noise records obtained in Hawaii during August of 1958 indicate an abnormal condition in the ionosphere which apparently persisted for most of the month, and which seems to be related to high altitude nuclear explosions over Johnston Island.

2. Observations

The system of noise recording used in the National Bureau of Standards network [1] [1] obtains a 15-min sample of the average noise power each hour on eight frequencies ranging from 13 kc to 20 Mc.[2] A median value of all observations made during a month at each hour of the day is determined at the end of each calendar month and is referred to as the "month-hour" median value of the noise. Necessary calibration factors are applied to convert to "db above ktb," this converted value being referred to as "F_a," or the effective antenna noise figure.

The Kekaha recording station is located on the island of Kauai, Hawaii. The nearest areas with high thunderstorm activity are on or near the major continental land masses several thousand miles distant. Records of weather stations on the islands of Kauai, Hawaii, and Oahu indicate a mean annual number of days with thunderstorms of less than ten, and no thunderstorms were observed at any of these stations during July or August 1958 [2]. It may be assumed, therefore, that the noise recorded at Kekaha in August was mainly noise originating at considerable distances from the station, and that most of it was propagated by the ionosphere.

From April through July of 1958, the median noise levels recorded at Kekaha indicated a remarkably constant pattern of noise: both the shape of the diurnal noise curve and the median observed levels of the noise power (for a particular time of day) showed only minor variations from month to month. The median levels for August, however, indicated some marked changes had occurred: from 51 kc to 2.5 Mc, the median nighttime levels were 5 to 15 db lower than previously, while the daytime levels at 51 kc and 160 kc were from 5 to 10 db higher than in previous months. The median diurnal range at 160 kc amounted to only 23 db in August, compared to 42 db in July; but in September the noise was once again very close to the July levels (fig. 1).

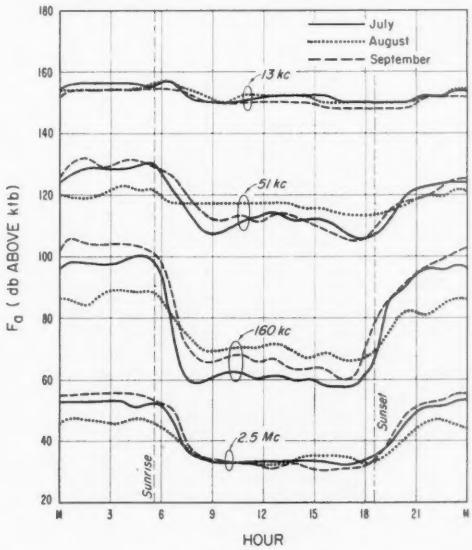


FIGURE 1. *Month-hour median values of radio noise at Kekaha.*

[1] Figures in brackets indicate the literature references at the end of this paper.
[2] Note: A 500-sec time constant is used in the averaging circuits.

The day-to-day and hour-to-hour variations in recorded noise power at Kekaha for the period July 25 to September 4, 1958, are shown in figures 2, 3, and 4. The period July 25 to 31 is indicative of normal conditions at Kekaha: the noise tends to reach its highest levels between 2200 and 0400 H.s.t. (Hawaiian standard time) with a fairly rapid fall starting at or shortly after local sunrise (about 0530 H.s.t. in late July) and reaching a minimum between 0800 to 1200 H.s.t. At 13 kc, however, the minimum value usually occurs in late afternoon; while at 51, 160, and 545 kc (and occasionally at higher frequencies) there are two minimums: about 0800 to 1000 and 1600 to 1800, with midday levels at these frequencies commonly 4 to 10 db above the minimum level.

Two nuclear devices were set off over Johnston Island (about 700 miles southwest of Kekaha) in August: one at an altitude of greater than 60 km on August 1 at 1050 G.c.t. (0050 H.s.t.), and a second at around 30 km on August 12 at 1030 G.c.t. (0030 H.s.t.) as announced by the AEC [3].

Referring to figure 2, note the abrupt drop in the received noise which occurred shortly after midnight on August 1, at a time of day when the noise is normally steady or rising.[3] The original recorder charts show that the drop in noise level was practically simultaneous with the explosion; the average noise power on the low frequency channel recording at the time (545 kc) began dropping rapidly at about 0050 H.s.t. and fell 18 db in 10 min. Because of the averaging time delay of the equipment, the actual decrease of the noise power during this interval was probably much greater, as indicated by the value obtained on 545 kc 1 hr later which was 32 db below the reading at 0050 H.s.t.

After the initial decrease in the noise, a partial recovery began 1 or 2 hr after the explosion, and it appears to have been nearly complete within a few hours at 5, 10, and 20 Mc as well as at 13 kc. However, at 51, 160, and 545 kc and at 2.5 Mc the nighttime levels were 10 to 30 db lower (than the preblast averages) for about six days, and conditions did not appear to be completely normal when the second explosion occurred on August 12.

Coincident with the second blast, the received noise level at Kekaha again dropped sharply, but the immediate effect of the explosion seemed to be somewhat smaller than after the August 1st blast. The long-term effects seem to have been somewhat greater, however, with a very slow rate of recovery at 51 kc to 2.5 Mc for 10 to 12 days afterward and a disturbed condition prevailing for about another 10 days. After about September 1, the recorded noise again seemed to be near normal (pre-August levels).

[3] Note: The apparent premidnight drop at 545 kc is attributable to the method of sampling used in the ARN–2 recorder. The value for the "00" hour is recorded at any time between 0000 and 0100 l.s.t., depending on the frequency; in this case, the values plotted at "00" for 545 kc and 20 Mc were actually recorded about 10 min after the blast, while the values plotted at "00" for all other frequencies were recorded before the blast.

FIGURE 2. *Radio noise power recorded at Kekaha, Hawaii: July 25 to August 7, 1958.*

Time of explosion indicated by arrows.

FIGURE 3. *Radio noise power recorded at Kekaha, Hawaii: August 8 to August 21, 1958.*
Time of explosion indicated by arrows.

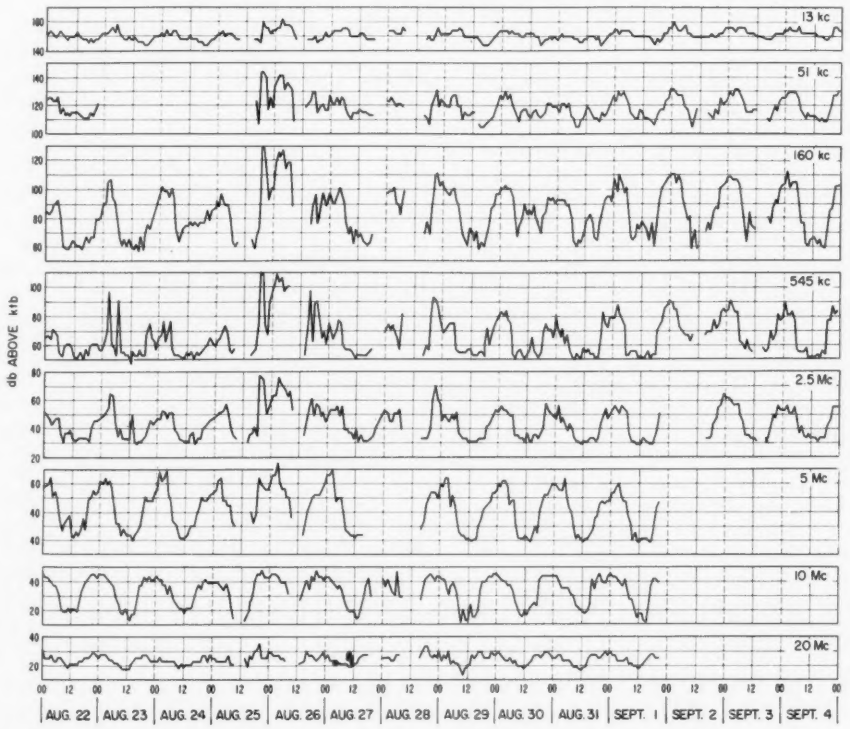


FIGURE 4. *Radio noise power recorded at Kekaha, Hawaii: August 22 to September 4, 1958.*

The unusually high levels recorded from 13 kc to 5 Mc on August 25 and 26 (fig. 4) may have been related to some local electrical phenomena on Kauai, since the station engineer reported a heavy overcast in the vicinity of Kekaha both days. However, no lightning or thunder were observed and no rainfall was recorded on either day in the southwest part of the island [4]. It seems more likely that some unusual ionospheric condition was responsible for the high noise levels recorded on these days.

As shown in figure 1, the median noise during the daytime at 51 kc and 160 kc was higher than normal during August. At 51 kc in particular, there was a "leveling out" of the received noise pattern in the daytime; the usual morning and afternoon minimums were much less pronounced or did not appear at all, and the overall daytime levels for several days after each explosion were close to the midday levels recorded during the pre-August period (fig. 2 and 3).

3. Conclusions

The records of the Kekaha noise station during August provide evidence of sudden and large magnitude changes in the ionosphere following the nuclear tests at Johnston Island. The main source of atmospheric noise at the time of day when the explosions occurred was probably in southeast Asia or Indonesia, resulting in a propagation path through or very near to the blast area. Apparently the region in which each explosion took place became very highly ionized, resulting in greatly increased absorption of signals normally propagated from levels of the ionosphere above that region.

Since an omnidirectional antenna is employed in the noise measurements, the horizontal extent of the abnormal ionization must have been rather large to produce changes in noise level of the degree observed at Kekaha. Some indication of the spread of ionization may be obtained from observations of signal strengths in a San Francisco, California to Hiraiso, Japan path at 13.75 Mc, which show a large loss of signal coinciding closely with the time of the first explosion [5]. Since this path is some 3600-km north of Johnston Island, a very rapid spread of ionization must have occurred to the north of Johnston. On August 12, however, when the explosion occurred at a much lower altitude, no significant effects were noticed on this circuit, although paths nearer Johnston were affected.

Because of the relatively short recombination time in the region where the explosions are thought to have occurred, the increase in absorption should have been of short duration. However, the Kekaha noise records indicate either a continued high level of absorption or a decrease in transmitted noise. The major sources of the noise received at Kekaha are believed to be several thousand miles distant, and it is extremely unlikely that overall thunderstorm activity in these regions could have been affected by the explosions to a sufficient degree to account for the changes in noise level which were observed. Therefore, it would seem that some mechanism in the ionosphere was acting to cause absorption of noise at night for an extended period at the frequencies from 51 kc to 2.5 Mc, while at the same time some improvement in daytime propagation may have occurred at 51 kc and 160 kc. Regardless of the exact mechanism involved, the Kekaha noise records provide evidence that high-altitude nuclear explosions may have a rather persistent effect on radio communications, particularly in the frequency range of 51 kc to 2.5 Mc.

Preliminary investigations of noise records from other stations in the NBS network have disclosed anomalies which may be related to the Johnston tests, but because of the more complicated nature of the received noise at these locations (compared to Kekaha), no conclusions can be reached until more thorough studies of the records can be made.

The author thanks W. H. Ahlbeck, station engineer at Kekaha, for the excellent series of noise measurements analyzed in this paper, and W. Q. Crichlow and W. F. Utlaut for helpful suggestions.

4. References

[1] W. Q. Crichlow, Noise investigation at VLF by the National Bureau of Standards, Proc. IRE, 45, 778 (1957).
[2] U.S. Weather Bureau, Local climatological data, 1958 (Lihue, Honolulu, Hilo).
[3] AEC Release, B–94 (June 15, 1959).
[4] Kekaha Sugar Co., Ltd., Rainfall records for August 1958.
[5] T. Obayashi, S. C. Coroniti, and E. T. Pierce, Geophysical effects of high-altitude nuclear explosions, Nature, 183, 1476 (1959).

Boulder, Colo. (Paper 64D1–35)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January-February 1960

Measured Frequency Spectra of Very-Low-Frequency Atmospherics[1]

Tatsuzo Obayashi[2]

(August 20, 1959)

New spectroscopes recording continuously the amplitude-frequency spectra of vlf atmospherics have been developed. Two receivers cover the frequency ranges 1 to 10 kc and 5 to 70 kc sweeping the respective bands repeatedly, and their outputs are displayed on intensity modulated cathode-ray tubes which are photographed on slowly moving film.

Observations have been carried out since June 1958, and it appears that the results provide an excellent experimental basis for comparisons with the mode theory of vlf ionospheric propagation. It is found that the frequency spectrum of distant atmospherics indicates a pronounced absorption near 3 to 5 kc, a broad intensity maximum around 10 to 20 kc, and a general decrease towards higher frequencies with undulating peaks. The selective absorption bands appearing in the spectrum are variable according to the time of day and seasons. These changes may be interpreted loosely as an ionospheric effect which is associated with the cutoff frequency of the waveguide bounded by the earth and the ionosphere. The solar flare effect on vlf atmospherics propagation is also revealed, which indicates a sudden shift of the spectrum to higher frequencies owing to the increase of ionization and the lowering of a reflecting height of the ionosphere.

1. Introduction

Considerable attention has been given recently to the very-long-distance propagation of radio waves in the very-low-frequency band (vlf). It is known that naturally-occurring atmospherics can provide excellent information about vlf propagation. Many investigators have found that vlf waves propagate to great distances with very small attenuation; however, they have also found evidence of a pronounced absorption band between 2 to 4 kc. Chapman and his colleague [1, 2],[3] for example, have made precise observations on waveform characteristics as well as the amplitude-frequency spectra of individual atmospherics. Their results showed that the spectrum of atmospherics varies with the distance of propagation, indicating the existence of an appropriate mode of propagation and of the strong selective attenuation due to ionospheric influences.

Meanwhile, a new theory of vlf ionospheric propagation, known as the mode theory, was suggested by Budden [3] in 1951. According to his theory, waves that have traveled considerable distances act as if they were propagated in the space between parallel reflecting surfaces representing the earth and the lower edge of the ionosphere, and the observed presence of an absorption band for frequencies of the order of 3 kc could be explained with this model when an appropriate value for the electron density of the reflecting layers is assumed. Further progress in theoretical studies have been made recently, particularly by Wait [4, 5], who has carried out elaborate computation of the modes of ionospheric propagation and showed that the characteristics predicted by the theory are in good accord with the experimental facts gathered from many sources.

In the present investigation, further experimental proof of the mode theory of vlf ionospheric propagation is presented. A new apparatus is described which records continuously the frequency spectrum of atmospherics. Two such radio spectroscopes covering the frequency ranges 1 to 10 kc and 5 to 70 kc are used. The basic action of the spectroscope is to scan the respective vlf bands continuously; the outputs are displayed on an intensity-modulated cathode-ray oscillograph and recorded photographically, instead of conventionally recording the amplitude of signals at several fixed frequencies. The pattern on the photographic film is, therefore, an intensity-modulated plot of frequency against time; and the variations of frequency spectrum such as the diurnal characteristics or ionospheric disturbances are clearly demonstrated.

It appears that the results provide excellent experimental proof for the mode theory of vlf ionospheric propagation. The results obtained so far are very encouraging. Some novel results are also found and their interpretations given in the present paper.

2. Measuring Equipment

The main purpose achieved with the present apparatus is the continuous observation and a direct display of the frequency spectra of atmospherics. Although there are several kinds of spectrum analyzers which could be applicable to the study of atmospherics, a scanning type of spectroscope is considered to be preferable because of the requirement of continuous observations.

The apparatus consists of two sets of radio spectroscopes covering the frequency ranges 1 to 10 kc and 5 to 70 kc. Each set has a conventional vlf atmospherics receiver, a frequency-scanning device, and a display unit including a photographic motion camera. A block diagram of the spectroscope is shown in figure 1. The antenna is a vertical whip with an effective height of 10 m. The receiver is a superheterodyne type consisting of hf and IF amplifiers,

[1] A preliminary account of this work is given by T. Obayashi, S. Fujii, and T. Kidokoro, An experimental proof of the mode theory of vlf ionospheric propagation, J. Geomagnetism and Geoelectricity **10**, 47 (1958).
[2] Radio Research Laboratories, Kokubunji, Tokyo, Japan (invited paper).
[3] Figures in brackets indicate the literature references at the end of this paper.

524214—59——4

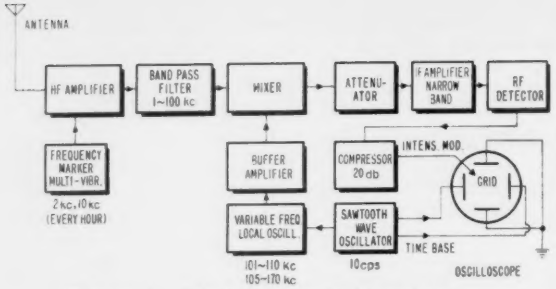


FIGURE 1. *A block diagram of vlf spectroscope.*

mixer, and detector. The local oscillator sweeps frequencies between 101 to 110 kc and between 105 to 170 kc by using a dust-core to alter its inductance. The sweep voltage is fed from a sawtooth generator. Simultaneously, a synchronized sweep is applied to the time-base of a cathode-ray oscilloscope so that a given horizontal position of the oscilloscope spot corresponds uniquely to a single frequency. By mixing this local oscillator with the external signal and applying a narrow-band IF transformer of 100 kc, the receiving frequency ranges are converted to 1 to 10 kc and 5 to 70 kc. The IF bandwidths are ± 600 cps respectively, and the frequency scanning rate is about 10 cps. These are appropriate for the present purpose. Special attention has been paid to provide a flat frequency response of the receiver over the swept-frequency range and with strong attenuation for the rest of frequencies. The overall frequency response of each receiver is shown in figure 2. It is evident that the receiver has a flat characteristic over the respective sweep-frequency range and a large attenuation more than 30 db at the rest of frequencies.

Thus, as the receiver sweeps the frequencies, giving complete covering of vlf band, the output is displayed as an intensity-modulated line on a cathode-ray oscilloscope. This is photographed on a slowly moving (1.5 cm/hr) 35-mm film, producing a frequency-time record with intensity of atmospherics appearing as variations in photographic density. The combination of the film characteristic and the logarithmic compressor permits the recording of a

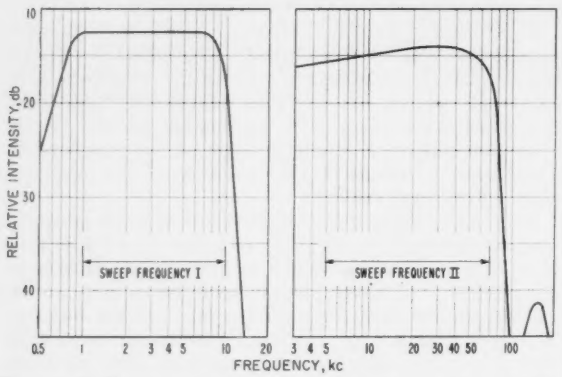


FIGURE 2. *Frequency response of the receivers. I and II.*

wide range of signal intensities (more than 20 db) in considerable detail. The record is calibrated by the frequency markers of the 2 kc- and 10 kc-multivibrators every hour which is switched on automatically by a standard time clock.

This type of vlf atmospherics spectroscope certainly has a major advantage over that of recording of atmospherics intensities at a number of single frequencies, and so far the results obtained have been very successful. Since the pattern on the film is an intensity-modulated plot, an auxiliary apparatus to reproduce an amplitude-frequency spectrum has also been designed. This is accomplished by successive scanning on the film using a photoelectric tube. An original amplitude of the signal is obtained in db scale with the aid of a suitable comparison technique.

3. Frequency Spectra of Atmospherics Propagated Through the Ionosphere

Observation of atmospherics using the present vlf spectroscopes has been carried out at the Hiraiso Radio Wave Observatory since June 1958. Simultaneously the continuous measurement of the intensity of atmospherics at 28 kc has also been made. Some of the records obtained are reproduced in figure 3, which are shown by the frequency-versus-local-time patterns, highlighted and dark portions indicating intense atmospherics and weak or no atmospherics, respectively. Signals from radio stations are also recorded, which appear at white straight lines indicated by cw. As can be seen clearly in these records, it is significant that there exists a strong absorption band for atmospherics propagation around 2 to 4 kc. The intensity of atmospherics is maximum at about 10 to 20 kc, and it decreases towards higher frequencies with remarkable undulating peaks. Considerable diurnal and seasonal variations are noticed.

In the discussion of these frequency spectra of atmospherics, however, the following point must be made in order to have an understanding of the frequency-versus-time characteristics of atmospherics displayed in the present records: since the duration of atmospherics is usually less than a few milliseconds, individual atmospherics photographed cover only a limited range of the frequency band, since the time-base of the oscilloscope is converted to the frequency sweep. In other words, the period of sweep frequency (order of 0.1 sec) is not short enough to cover the whole spectrum of individual atmospherics. However, if it is assumed that numerous atmospherics are occurring at random within a short duration, the film recorded with a sufficiently long time exposure should reveal the relative amplitude-frequency spectrum in a loose statistical sense.

Looking through the records, it has been found that the short-distance atmospherics (probably less than 100 km) do not show complicated spectral patterns, but they indicate rather a flat spectrum over wide frequency range (fig. 4). Therefore, the

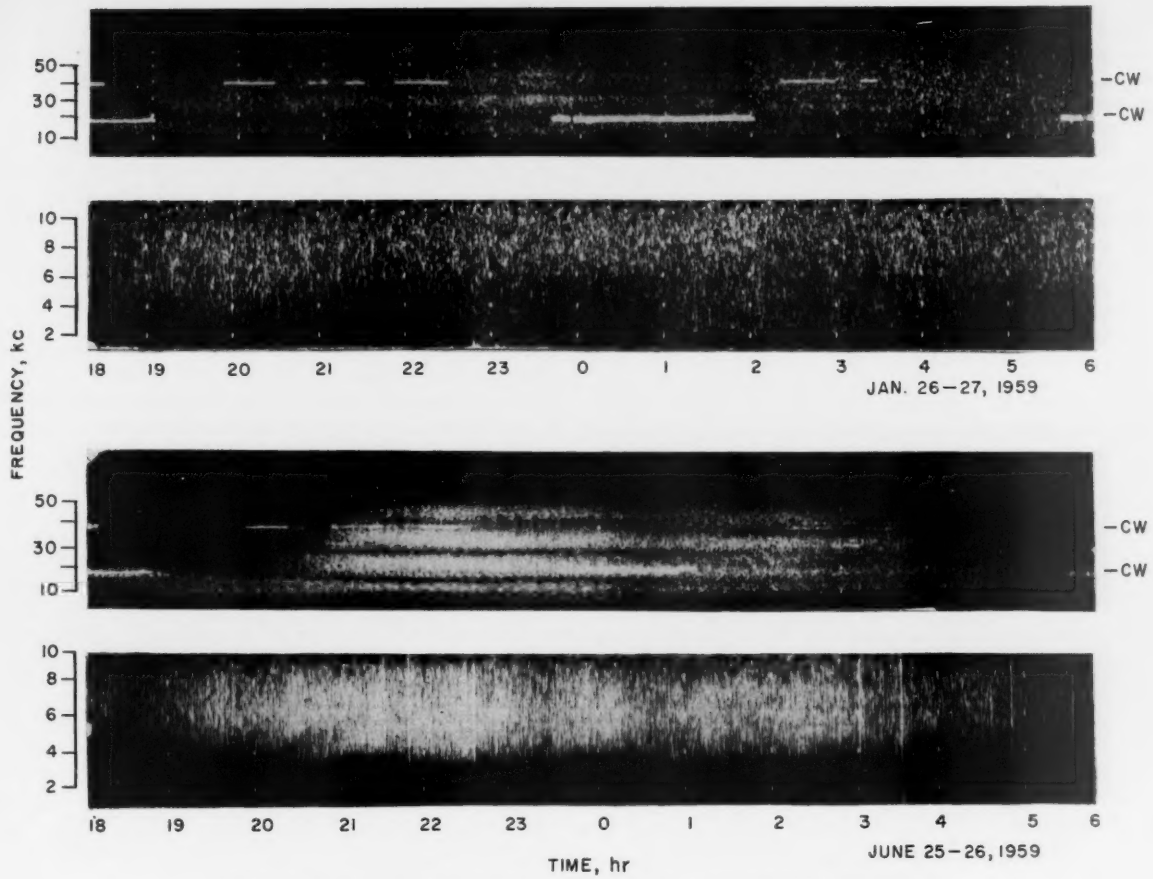


FIGURE 3. *Frequency spectra of vlf atmospherics of ionospheric propagation.*

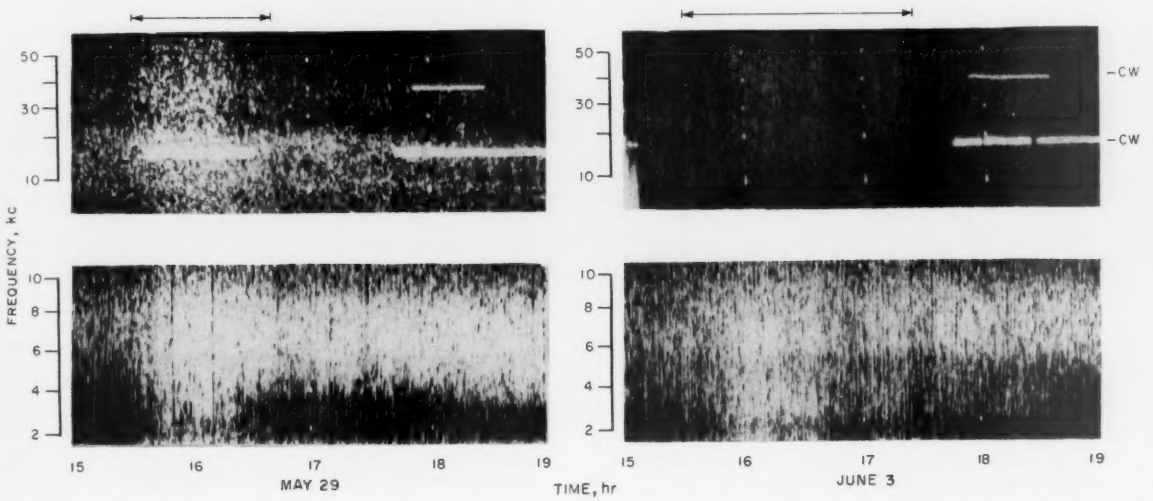


FIGURE 4. *Records of short-distance atmospheric, indicating a flat spectrum over a wide frequency range (↔ time during which local thunderstorms were active).*

outstanding features of strong selective absorption bands appearing in the records must be due to atmospherics propagated through considerable distances, which may be largely influenced by the ionospheric conditions.

In order to show more clearly the observed characteristics of vlf atmospherics, detailed analysis is made using those data of measured amplitude frequency spectra, which are obtained with the aid of the photoelectric scanner. In figure 5, average diurnal variations of amplitude-frequency spectrum of atmospherics for winter (January) and summer (June) seasons are shown, in which local short-distance thunderstorms are excluded. The characteristics of vlf radio-wave propagation are well represented, though the detailed structure of spectrum is rather smoothed out. For comparison, the average diurnal variations at the fixed frequency, 28 kc, are also reproduced in figure 6. Double maximums appear in the summer afternoon. The one around 16 hr is, however, more or less attributed to the thunderstorms in the vicinity of Japan, and hence the spectrum of atmospherics during this time of day shows a slightly different nature when compared with other seasons. From this analysis, it is confirmed that the nighttime spectrum shows strong absorption at about 2 to 3 kc and has a broad maximum at 5 to 50 kc with several undulating peaks. On the other hand, the daytime spectrum has smooth double maximums at 15 kc and 40 kc, its intensity being weaker than that of nighttime. Such general characteristics are well illustrated in figure 7, in which typical amplitude-frequency spectra at nighttime and daytime are shown.

One of the most interesting phenomena appearing in the records is the effect at the transition of spectrum between daytime and nighttime. Since two

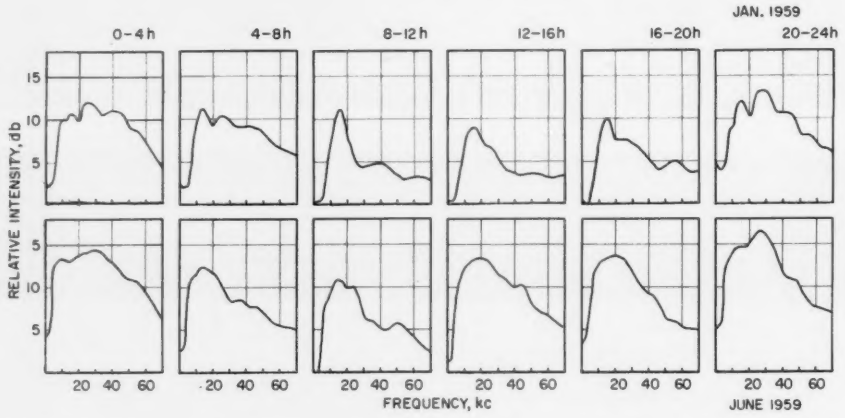


FIGURE 5. *Average diurnal variations of vlf amplitude-frequency spectrum at atmospherics in winter and summer seasons (shown in upper and lower sets, respectively).*

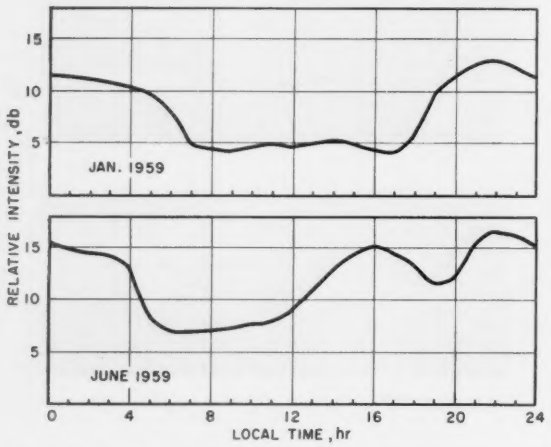


FIGURE 6. *Average diurnal variations of integrated field intensity of atmospherics (28 kc) observed at Hiraiso.*

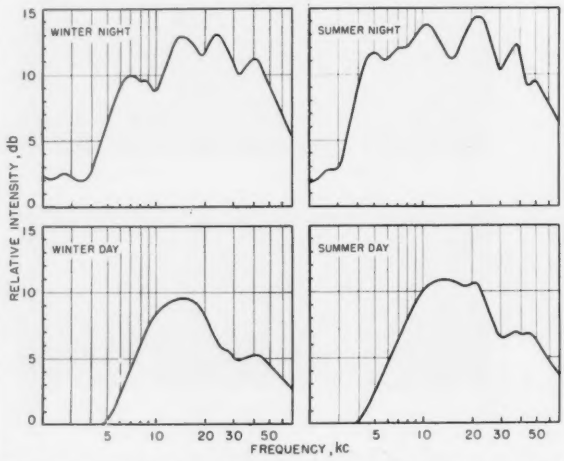


FIGURE 7. *Typical amplitude-frequency spectra of distant vlf atmospherics for nighttime and daytime.*

types of spectrum interchange rather distinctly at local sunset, the daytime maximum about 15 kc is shifted towards lower frequencies, while the effect at sunrise is reverse. Changes in cutoff frequency and the shift of spectral pattern at local sunset or sunrise are shown by several records in figure 8.

These characteristic features of the amplitude-frequency spectrum of vlf atmospherics cannot be explained by the simple Austin-Cohen law of vlf propagation, but more likely by the waveguide mode theory. The conception that vlf radio waves propagate to great distances via multiple reflections in the waveguide bounded by the earth and the ionosphere has been developed by Budden [3] and Wait [4] as the mode theory of vlf ionospheric propagation. An essential part of this theory is the existence of definite modes of propagation. In order to simplify the evaluation of these mode characteristics, let us consider the ground and the ionosphere as a waveguide of perfectly conducting plates. Then, the cutoff wavelengths would be given by $2h/n$, where n is the integer specifying the number of the mode and h is an effective reflecting height of the ionosphere. When the boundary of the waveguide is not perfectly conducting, as the case of the actual ionosphere, the cutoff wavelength is not sharply defined and, crudely speaking, the effective reflecting height of the ionosphere appears to be higher [4].

From the above consideration, it may be expected that the effect of the transition of spectrum from daytime to night conditions, which appeared as the shifting of the maximum intensity and also the cutoff frequencies towards lower frequencies, can be explained as the upward shift of the ionospheric height or else the decrease of the conductivity in the lower ionosphere. On the other hand, it is already well known from a number of experiments that the reflecting height of vlf radio waves in the ionosphere varies consistently from 70 km during the day to 90 km at night [6, 7]. Therefore, this is in accord with the present conclusion.

There is another interesting fact lending further support to the mode theory of vlf propagation. That is the effect of SID's, which are known as the sudden enhancement of atmospherics (SEA). Many SEA's accompanied by prominent solar flares were observed. Two such examples of changes of the spectrum of atmospherics are shown in figure 9 with the records of integrated level of atmospherics at 28 kc. A solar flare occurred at 16 hr 50 min (l.t.), March 29, 1959, which was associated with a very pronounced SEA. A sudden shift of the frequency spectrum at the beginning of the solar flare is evident. As shown in figure 10, the band of strong intensity of atmospherics was around 10 to 15 kc before the flare as in the case of usual daytime, and then at

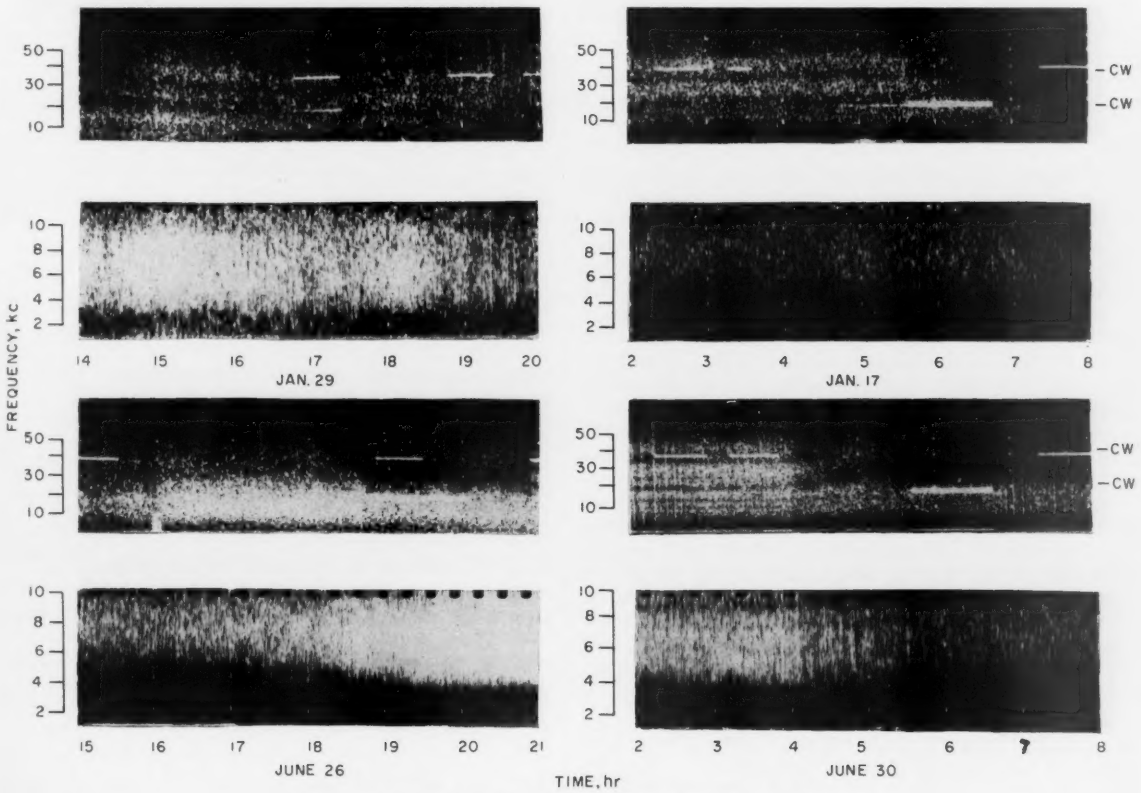


FIGURE 8. *Changes of spectrum at local sunset or sunrise.*

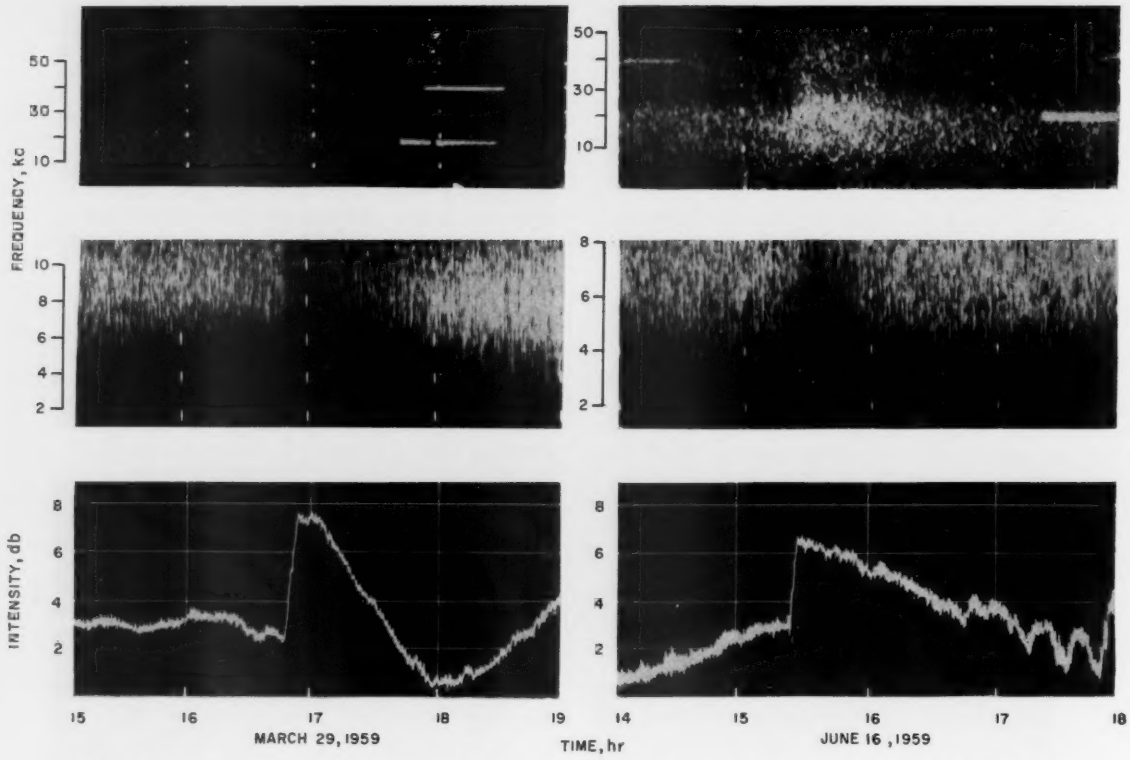


FIGURE 9. *Changes of spectrum during SEA's and records in the integrated intensity of atmospherics at 28 kc.*

16 hr 50 min the intensity was raised up, shifting the band of strong intensity towards the higher frequency band of vlf. The intensity of distant atmospherics is also enhanced, while at the frequency below 10 kc a sudden reduction in intensity occurs. Another record shows the effect of a solar flare at 15 hr 25 min (l.t.), June 16, 1959. This SEA occurred during rather greater thunder activity; however, the same effect mentioned above can be seen considerably well. It has been confirmed that all SEA's found in the records of 28 kc showed such a sudden shift of spectral pattern towards higher frequencies, though the amount of shift was variable for individual cases.

Since it is known that during SID's the electron density in the lower ionosphere is increased and its height is lowered considerably, these facts would eventually yield the explanation that the reduction of field intensity in the lower vlf band is caused by the shifting of the cutoff towards slightly higher frequencies, while in the frequency above the cutoff the intensity is enhanced owing to the good reflection condition in the lower edge of the ionosphere.

The observed stripe of absorption bands is still puzzling. However, it is expected that the actual observed intensity of distant atmospherics must be the sum of contributions from many waveguide modes [4], which are appreciably influenced by the distance from the source and the condition of the lower ionosphere. A detailed investigation of the theory of the waveguide modes would possibly give the appropriate solution.

4. Discussion of the Results

In the discussion of the frequency spectrum of atmospherics, the waveform characteristics must be mentioned, which have been studied extensively by many researchers in this field. Since the investigation of waveforms and of frequency spectrum are merely different approaches to a single phenomenon, their results must be mutually equivalent. Mathematically, the waveform $G(t)$, the amplitude-frequency spectrum $S(\omega)$ and the phase-frequency spectrum $\phi(\omega)$ have the following mutual relations:

$$G(t)=\frac{1}{\pi}\int_0^{\infty} S(\omega)\cos\left(\omega t+\phi(\omega)\right)d\omega$$

$$S(\omega)=\sqrt{\left(\int_{-\infty}^{+\infty} G(t)\cos\omega t\,dt\right)^2+\left(\int_{-\infty}^{+\infty} G(t)\sin\omega t\,dt\right)^2}$$

and

$$\phi(\omega)=\tan^{-1}\left(-\int_{-\infty}^{+\infty} G(t)\sin\omega t\,dt\Big/\int_{-\infty}^{+\infty} G(t)\cos\omega t\,dt\right)$$

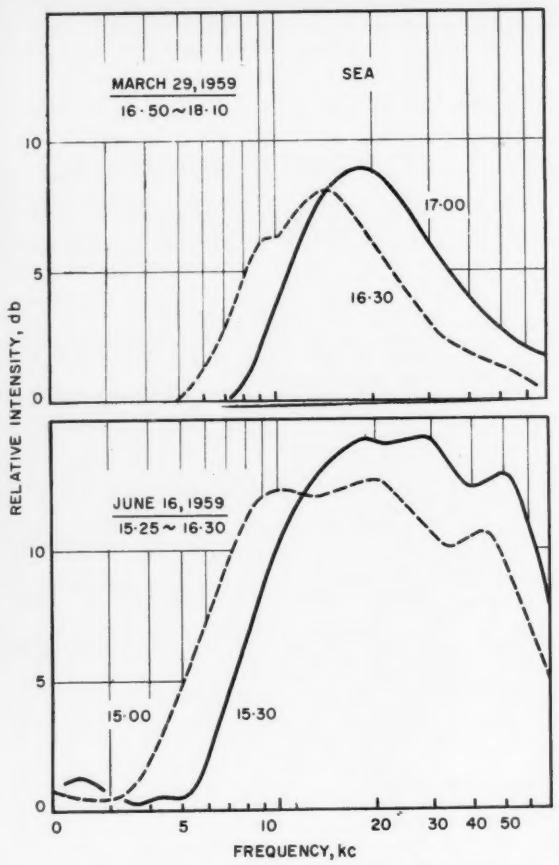


FIGURE 10. *Change of frequency spectrum at the SEA on March 29 and June 16, 1959 (---- presolar flare; —— during SEA.)*

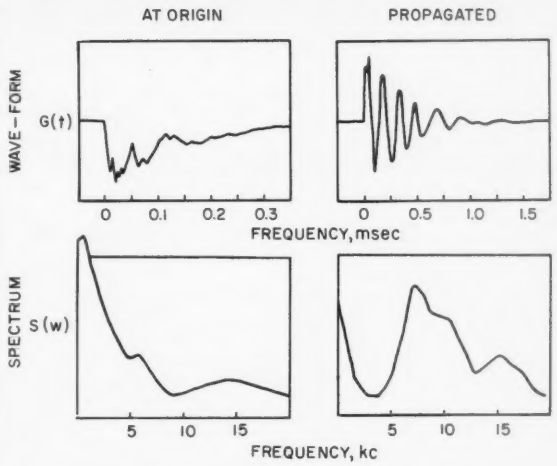


FIGURE 11. *Waveforms of atmospherics and their frequency spectra (after K. Sao [8]).*

Therefore, from the known waveform it is possible to compute its frequency spectrum. Using these Fourier transform expressions, Sao [8] has made the frequency analyses of a number of waveforms of atmospherics. Two types of frequency spectrum are characteristic; one is of the origin of an atmospheric and the other is of a distant atmospheric. Those waveforms and their spectra obtained by him are reproduced in figure 11. The frequency spectrum at the origin of a lightning stroke shows a gradual decrease of amplitude towards higher frequencies, having almost flat characteristics at about 10 to 30 kc. This result is in accord with the frequency spectrum of short-distance atmospherics shown in figure 4. Waveforms of distant atmospherics are more complicated and show the influences of the ionosphere. They often have the appearance of a damped sinusoid; the apparent frequency within the pulse is decreasing with time whereas the mean spectral maximum is increasing with range. Such a behavior has been predicted theoretically by Wait [5, 9]. The derived spectrum indicated heavy attenuation around 2 to 5 kc and small attenuation at around 10 kc, which is generally in good agreement with our observations.

Another investigation has been carried out by Chapman and Matthews [1] who used a spectroscope consisting of a number of narrow band-pass filters tuned to frequencies in the range of 40 cps to 16 kc, recording the amplitude-frequency spectra and corresponding waveforms for individual atmospherics. They also discussed the relation between the waveform characteristics and the significant changes in the spectrum as the distance of propagation increases. Summing up their results, Chapman and Macario [2] derived attenuation curves with respect to frequency for those of distant atmospherics. It is indicated that in the frequency range from 200 cps to 10 kc the greatest increase of attenuation occurs at about 2 kc, it being about 20 db per 1,000 km by day and 10 db by night.

Here, it might be worthwhile to link up this with our result. For this purpose, the frequency spectra shown in figure 7 can be used. Since their curves were given by attenuation in decibels per 1,000 km, our results must be converted to this scale. Although the distribution of the originating source of atmospherics cannot be identified in the present observation, according to Kimpara and Kimura [10] distant atmospherics received in Japan mainly come from regions in south and east China and the Philippines. Therefore, the average distance from the sources is of the order of 2,000 km. Assuming this, and also that the minimum attenuation around 10 kc is about 2 to 3 db per 1,000 km, the curve of the variation of attenuation with frequency in the range 1 to 100 kc is illustrated in figure 12. This curve, for the frequency range of 1 to 10 kc, can be compared with Chapman and Macario's. Also Taylor and Lange [11] reported nighttime attenuation values in good agreement with values presented here. These results

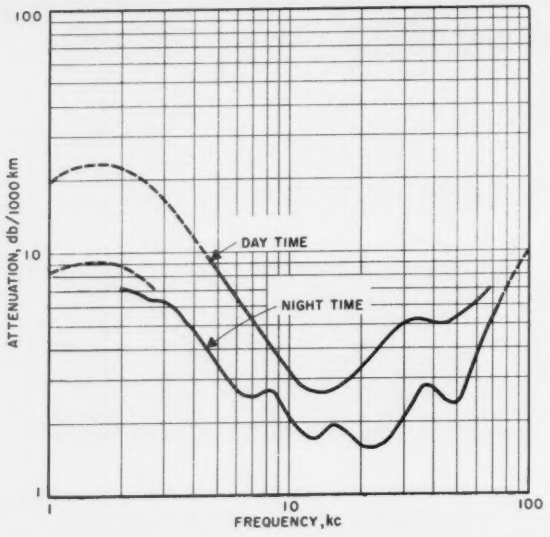


FIGURE 12. *The variation of attenuation with frequency.*
The dotted curves for 1- to 10-kc range are taken from Chapman and Macario [12].

give a general idea of the way vlf waves are attenuated and can propagate to great distances.

In concluding, it is emphasized that the continuous observation of the frequency spectrum of atmospherics reveals important characteristics of vlf radio waves propagated through the ionosphere. The general agreement between the experimental and theoretical results substantiates the mode theory. The fact that certain bands of vlf radio waves propagate to great distances with small attenuation accounts for the success of long-range navigational systems, worldwide communications or standard radio-wave systems, and the tracking of atmospheric storms.

5. References

[1] F. W. Chapman and W. D. Matthews, Audio-frequency atmospherics, Nature **172,** 495 (1953).

[2] F. W. Chapman and R. C. V. Macario, Propagation of audio-frequency radio waves to great distances, Nature **177,** 930 (1956).

[3] K. G. Budden, The propagation of a radio atmospheric, Phil. Mag. **42,** 1 (1951); **43,** 1179 (1952).

[4] J. R. Wait, The mode theory of vlf ionospheric propagation for finite ground conductivity, Proc. I.R.E. **45,** 760 (1957).

[5] J. R. Wait, The attenuation vs. frequency characteristics of vlf radio waves, Proc. I.R.E. **45,** 768 (1957).

[6] A. H. Waynick, The present state of knowledge concerning the lower ionosphere, Proc. I.R.E. **45,** 741 (1957).

[7] J. R. Wait, Diurnal change of ionospheric heights deduced from phase velocity measurements at vlf, Proc. I.R.E. **47** (1959).

[8] K. Sao, Researches in the frequency analyses of waveforms of atmospherics, Proc. Research Inst. Atmospherics, Nagoya Univ. **5,** 12 (1958).

[9] J. R. Wait, Propagation of very-low-frequency pulses to great distances, J. Research NBS **61,** 187 (1958) RP 2898.

[10] A. Kimpara and Y. Kimura, On field intensity recording of atmospherics at 27 kc, Proc. Research Inst. Atmospherics, Nagoya Univ. **5,** 21 (1958).

[11] W. L. Taylor and L. J. Lange. Some characteristics of vlf propagation using atmospheric waveforms, Recent advances in atmospheric electricity (Pergamon Press, New York, N.Y., 1959).

BOULDER, COLO. (Paper 64D1–36)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Determination of the Amplitude-Probability Distribution of Atmospheric Radio Noise From Statistical Moments

W. Q. Crichlow, C. J. Roubique, A. D. Spaulding, and W. M. Beery

(July 23, 1959)

During the International Geophysical Year, the National Bureau of Standards established a network of atmospheric noise recording stations throughout the world. The ARN–2 noise recorder at these stations measures three statistical moments of the noise: average power, average voltage, and average logarithm of the voltage. An empirically-derived graphical method of obtaining an amplitude-probability distribution from these three moments, and its development, is presented. Possible errors, and their magnitudes, are discussed.

1. Introduction

The interference to reception of radio signals caused by atmospheric noise depends not only on its average level but also on its detailed characteristics. A complete description of the noise received at a particular location would require an exact determination of the variation of its instantaneous amplitude as a function of time. Because of the complexity of the noise structure and the fact that no two samples of noise have identical amplitude-time functions, it is necessary to resort to simpler types of description of a statistical nature.

Various methods of measurement have been investigated by experimenters throughout the world for many years, and the relationship between the measured values and the interference caused to various types of service has been studied.

In order to determine the most significant characteristics to be measured, an international subcommittee was formed within Commission IV of URSI [1] to study the problem. A recommendation [1],[2] based on the subcommittee's report [2], was endorsed by Commission IV at the XIIth General Assembly of the URSI in 1957. In the Annex to this recommendation, it is stated that "it would be highly desirable to obtain detailed statistical information on the amplitude-probability distribution of the instantaneous envelope voltage, the various time functions, and the direction of arrival of the noise at many locations throughout the world. However, since continuous detailed measurements of this type at all frequencies and a large number of stations become prohibitive because of the complexity of the equipment involved and the large number of personnel necessary to carry out the observations, it is recommended that these complete studies of the detailed noise characteristics be confined to a few selected locations with continuous measurements of one or more simple parameters at a large number of stations."

[1] International Scientific Radio Union.
[2] Figures in brackets indicate the literature references at the end of this paper.

During the International Geophysical Year, the National Bureau of Standards established a network of 16 noise recording stations distributed throughout the world [3]. Measurements are made with the ARN–2 recorder which was designed at NBS to record the average noise power in a bandwidth of about 200 cps on eight fixed frequencies between 13 kc and 20 Mc. In order to obtain additional information on the character of the noise, two other statistical moments, the average envelope voltage, and the average logarithm of the envelope voltage are recorded at ten of the stations. These data are available from IGY World Data Center A, NBS, Boulder, Colo., at a nominal fee. The results of data analysis will also be published in an NBS Technical Note for public sale.

The amplitude-probability distribution is very useful in assessing the interference potential of noise and has been measured at several locations [4, 5, 6, 7]. Since data on the moments measured by the ARN–2 can be recorded continuously over the full frequency range with much less equipment and fewer personnel than would be required for a comparable amount of data on the complete amplitude-probability distribution, an investigation was made at the NBS of methods of deriving the complete distribution from these three moments.

It is the opinion of the authors that mathematical models for the distribution [1,5,7,8] have proved to be extremely cumbersome and of doubtful value in deriving the distribution from measured moments; therefore, an empirically derived graphical method has been developed.

2. Measurements of the Amplitude-Probability Distribution

Simultaneous measurements of the amplitude-probability distribution and the three moments measured by the ARN–2 were needed over a wide frequency range with a variety of types of noise. There were two specific purposes in making these measurements:

a. The distribution measurements, alone, were needed to establish typical shapes that could be expressed by numerical parameters, easily related to the measured statistical moments.

b. The simultaneous measurements of the distributions and the moments were needed to check the accuracy of the ARN–2 noise recorder in measuring the moments.

Accordingly, a distribution meter was constructed to be used in conjunction with the ARN–2 recorder.

The operation of the distribution meter is based on counting techniques. The total number of IF cycles in the ARN–2 which exceed a specified amplitude in a given time interval are counted. By proper choice of the time interval, the percentage of time a level is exceeded can be read directly from the counter. Three counters are used to provide simultaneous measurement at three levels, separated in amplitude by 6 db. By use of a calibrated, manually-operated attenuator, the total possible range of 100 db can be covered in levels separated by multiples of 2 db. The time duration of each sample can be preset to provide an automatically timed sample having an interval ranging from 1 sec to 1 hr, in 1 sec increments.

The actual intervals were chosen of sufficient duration to provide a statistically significant sample of a slowly varying phenomenon such as thunderstorm activity. Longer intervals were required for the high-level, low-probability noise pulses than for the low-level, high-probability noise. The time necessary to obtain a complete curve ranged from about 15 min to 30 min. The longer period was necessary for measuring noise with a large dynamic range.

A check for accuracy of operation was made by obtaining a distribution of thermal noise. The Rayleigh distribution thus obtained confirmed the accuracy of measurement. Figure 1 is a typical amplitude-probability distribution of atmospheric radio noise at 13.3 kc. This was plotted on Rayleigh graph paper, whose coordinates were chosen so that a Rayleigh distribution will plot as a straight line with a slope of $-\frac{1}{2}$.[3] The coordinates are shown as noise level in decibels above the root mean square voltage versus the percentage of the time that each level is exceeded. This measurement was made with the distribution meter at the Boulder Laboratories. The dynamic range between the 0.0001 percent and 99 percent intercepts is approximately 84 db.

It was found that these amplitude-probability distributions for atmospheric radio noise can be adequately represented by a three-section curve which has been drawn through the data points in figure 1. This curve has a characteristic shape that can be described by means of four numerical parameters to be defined later. The accuracy of fit using this particular kind of curve composed of three sections is typical of more than 100 distributions experimentally measured at Boulder. In the particular case of thermal noise, this curve becomes a straight line.

[3] These coordinates are log of voltage versus $-\frac{1}{2}\log_{10}(-\ln \text{ of probability})$.

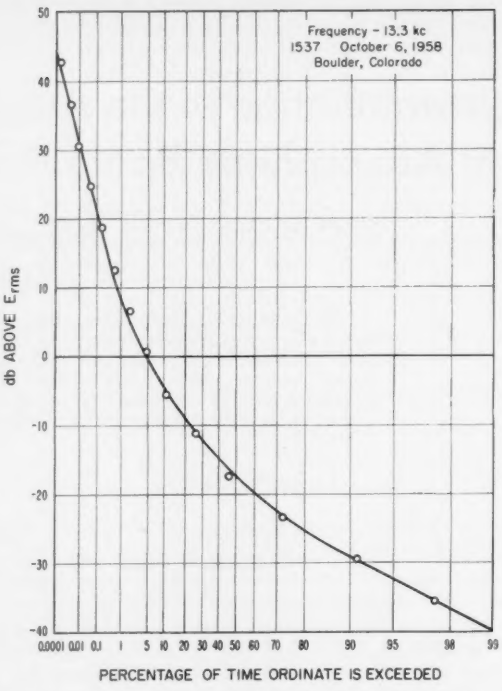


FIGURE 1. *Measured amplitude-probability distribution of atmospheric radio noise.*

Data obtained by other experimenters [2] in Alaska, Panama, England, and Florida when plotted on Rayleigh paper also fit curves of this form.

The lower portion of the curve, representing low voltages and high probabilities, is composed of many random overlapping events, each containing only a small portion of the total energy. Therefore, this portion of the curve must approach a Rayleigh distribution.

The section representing very high voltages exceeded with low probabilities is, in general, composed of nonoverlapping large pulses occurring infrequently. From experimental measurements of atmospheric noise distributions, this section has been found to be well represented by a straight line on Rayleigh graph paper. On this graph paper, the remaining section of the distribution has been found to correspond quite closely to an arc of a circle tangent to the above two straight lines. Based on experimental evidence, it is assumed that no pulses of amplitude greater than those for which the probability of being exceeded is 10^{-6} are present in the distribution.

3. Definition of Parameters and Statistical Moments

Since on the coordinate paper used for the plotting of amplitude distributions for atmospheric radio noise a Rayleigh distribution will always have a slope of $-\frac{1}{2}$, and since the center of the circular arc men-

tioned above will lie on the bisector of the angle formed by the Rayleigh and high-voltage, low-probability lines, four parameters are needed to define the distribution. A point through which the Rayleigh line passes, a point and a slope for the high-voltage, low-probability line, and a parameter of some kind to determine the radius of the circular arc are needed.

These four parameters are defined as follows:

a=point at which the Rayleigh line intersects the 0.5 probability coordinate.

b=point on that tangent to the circular arc which is perpendicular to the bisector of the angle formed by the Rayleigh and high-voltage, low-probability lines and on the probability coordinate through the vertex of this angle.

c=point where the high-voltage, low-probability line intersects the 0.01 probability coordinate.

s=slope of the high-voltage, low-probability line.

The three statistical moments experimentally determined are defined as follows:

$$\text{rms voltage: } e_{\text{rms}}=\sqrt{\int_0^1 e^2 dp}$$

$$\text{average voltage: } e_{\text{ave}}=\int_0^1 e\,dp$$

antilog of the mean log of voltage:

$$e_{\log}=\text{antilog}\int_0^1 \log e\,dp.$$

where e is $1/\sqrt{2}$ times the instantaneous envelope voltage, and p is the probability of e being exceeded.

Since the ARN–2 noise recorder measures the db difference betweeen the e_{rms} and e_{ave} and between the e_{rms} and $e_{\log}$, and in order to minimize the necessary graphical calculations, the above parameters are altered to relate them to the Rayleigh section of the distribution as follows (see fig. 2):

$A=20\log\frac{a}{e_{\text{rms}}}$=db difference between Rayleigh at 0.5 probability and the e_{rms} value.

B=db difference between the point of intersection of the Rayleigh and high-voltage, low-probability lines and the tangent to the circular arc. (The tangent is drawn perpendicular to the bisector of the angle formed by the Rayleigh and high-voltage, low-probability lines.)

C=db difference between c and value of Rayleigh at the 0.01 probability coordinate.

X=absolute value of slope of high-voltage, low-probability line relative to Rayleigh.

$X=-2s$

The moments actually measured are labeled as follows:

$$V_d=20\log\frac{e_{\text{rms}}}{e_{\text{ave}}}$$

$$L_d=20\log\frac{e_{\text{rms}}}{e_{\log}}$$

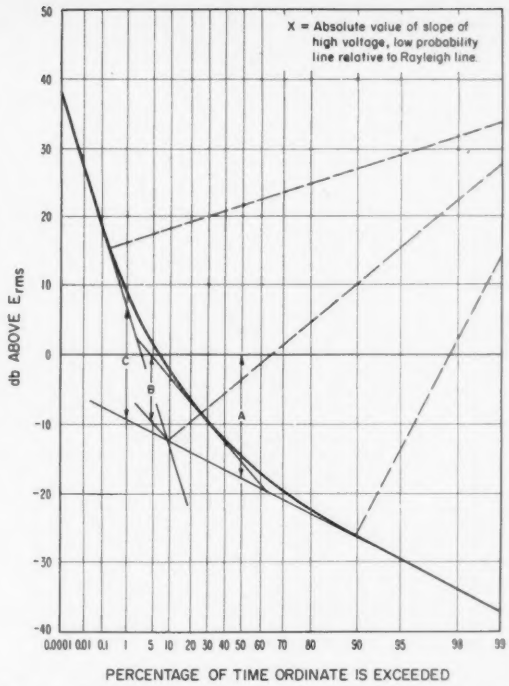


FIGURE 2. *Definition of parameters for amplitude-probability distribution atmospheric radio noise.*

The e_{rms} is also measured.

The following relation will always hold true for the distribution:

$$e_{\log}<e_{\text{ave}}<e_{\text{rms}}$$

therefore, V_d and L_d are always positive.

4. Determination of the Parameters as Functions of the Measured Moments

Since four parameters are needed and only three statistical moments are measured, a dependency between two of the parameters must be determined in order to be able to obtain the remaining three independent parameters as functions of the measured moments. From experimentally measured distributions at eight frequencies and various bandwidths, the following linear relation between X and B is found to hold in general (fig. 3).

$$B=1.5(X-1)$$

Distributions were employed in which C was allowed to vary from 0 to 40 db and X from 1 to 12. These distributions were numerically integrated in order to determine their three moments. From these integrations, V_d and L_d were obtained as functions of X and C. A was also obtained as a function of X and C (fig. 4). These functional relations were solved simultaneously by graphical means to obtain X and C as functions of V_d and L_d (figs. 5 and 6).

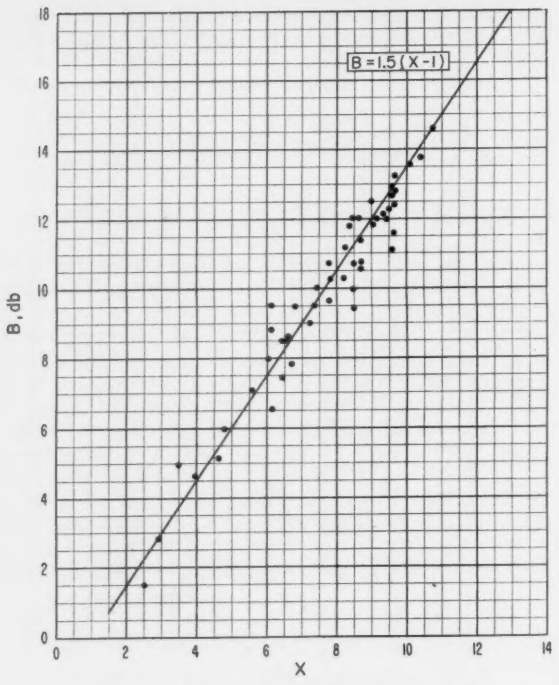


FIGURE 3. *Experimental correlation of* B *and* X.

Figure 7 shows the distribution obtained by the above procedure for the following measurements:

$$V_d = 4.5 \text{ db below } e_{rms}$$

$$L_d = 9.2 \text{ db below } e_{rms}.$$

On the X versus V_d and L_d curves, the above measurements are seen to give an X of 3.7, and on the C versus V_d and L_d curves the above measurements give a C of 15.8 db. Using the A versus X and C curves, this pair of values for X and C gives an A of 12.3 db. Point a is located on the 50 percent coordinate 12.3 db below the E_{rms} and the Rayleigh line with slope $-\frac{1}{2}$ is drawn. Point c is located 16 db above the point where the Rayleigh line intersects the 1 percent coordinate, and through point c the high-voltage, low-probability line with slope $s = -X/2 = -1.8$ is drawn. The angle formed by the Rayleigh line and the high-voltage, low-probability line is now bisected and B is calculated from the relationship $B = 1.5(X-1)$. B is found to be 3.6 db. Point b is located 3.6 db above the intersection of the Rayleigh and high-voltage, low-probability lines and through this point the tangent line is drawn perpendicular to the angle bisector. Distance 2 to 3 is made equal to distance 1 to 2, and at point 3 on the Rayleigh line a perpendicular to the Rayleigh line is constructed. This perpendicular intersects the angle bisector at the center of the circular-arc portion of the distribution. The complete distribution is now determined with the construction of this arc.

It should be noted from the curves (figs. 5 and 6) that for a given V_d only a certain range of values of L_d will result in a distribution of the above form. These ranges are enclosed approximately within the dashed lines. Also, the combination of the above given V_d and the minimum allowable L_d results in a nonunique solution, i.e., there is, essentially (within the range of accuracy of measurements), an infinite number of distributions, all with the same V_d and L_d of the above special combination. Fortunately these special combinations almost never occur.

The above facts allow us to determine to some extent the validity of data received from recording stations, and whether a measured sample is true atmospheric noise or contaminated noise.

5. Errors in Noise Measurements and Their Influence on the Calculated Distribution

The accuracy with which a distribution can be determined from measured moments depends upon (1) the validity of the form factor which has been taken to represent the distribution, and (2) the accuracy of the measured statistical moments.

The validity of the form factor is very difficult to check because of the fact that 15 to 30 min are required to measure the distribution, and the statistics of the noise do not always remain constant for this period of time. As a check for changing noise characteristics, each series of distribution measurements was immediately repeated and any changes noted. Further, the moments recorded by the ARN–2 provided information on the stability of the noise statistics. Figure 1 represents an example of the fit in the case where the noise moments were measured precisely and the noise characteristics appeared to be constant during the interval required for the measurement of the entire distribution. Methods of recording noise samples on magnetic tape are under development and these samples will enable a more critical study of the form factor as well as other detailed characteristics.

Errors in the ARN–2 could be detected by comparing the recorded three parameters with the same three parameters as numerically integrated from the curve recorded by the distribution meter. The most likely source of error in the ARN–2 was the possible saturation of the square law detector with a resultant lowered value for E_{rms}. Since the average voltage and average logarithm of the voltage are measured in decibels relative to E_{rms}, their values would be in error by the same amount.

Approximately sixty sets of distribution measurements were analyzed for the error made by the ARN–2 recorder in measuring the moments. These ARN–2 errors were divided into class-intervals and plotted as error-probability curves on arithmetic probability graph paper in figure 8. The fairly straight lines obtained indicated largely random errors, with some bias due to square-law detector saturation. For the E_{rms} values, the standard deviation

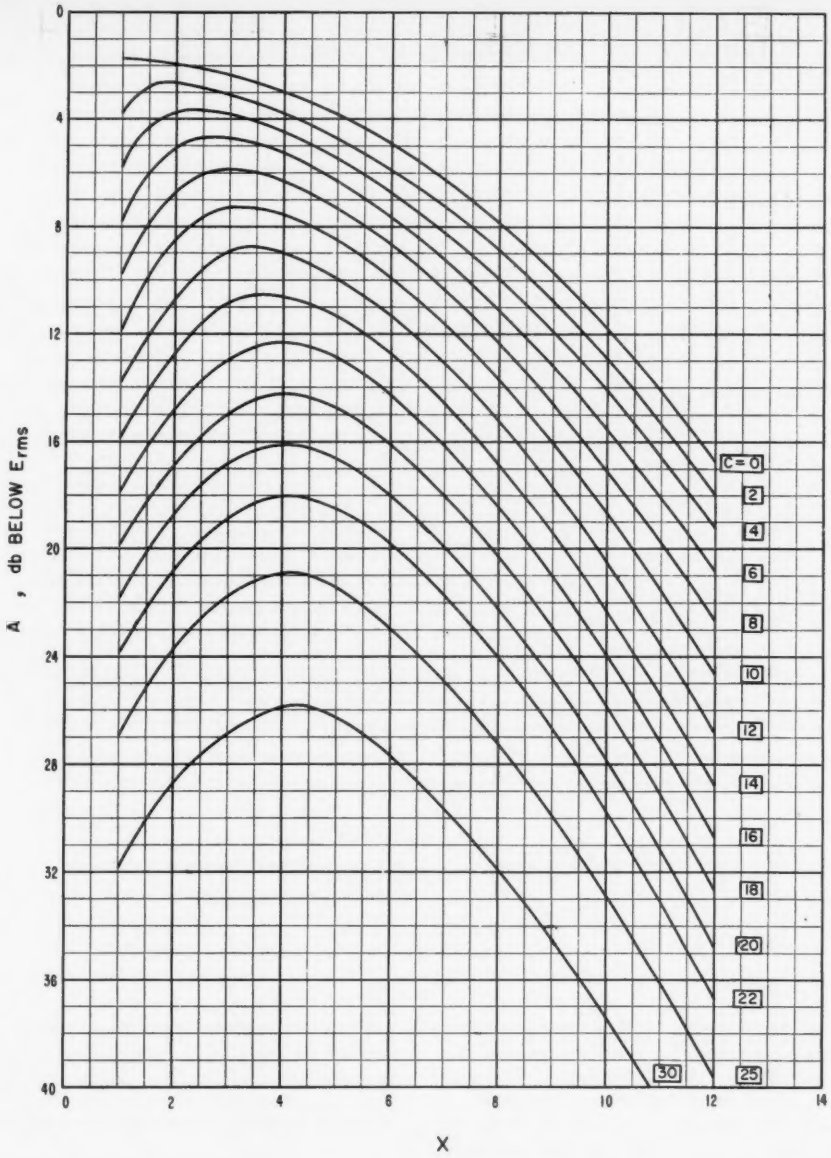


FIGURE 4. A *versus* X *and* C.

was about 1.1 db, with an average error of −0.4 db. Ninety percent of the values were within ±2 db of the mean. For the E_{ave} values, the standard deviation was about 0.8 db, with an average error of +0.2 db. Ninety percent of the values were within ±1.5 db of the mean. For the E_{log} values, the standard deviation was about 1.4 db, with an average error of +0.6 db. Ninety percent of the values were within ±2.5 db of the mean.

An effort was made to correlate the error with dynamic range between the 0.0001 percent and 99 percent intercepts of the amplitude-probability distributions. The same data were divided into class-intervals and the medians of the class-intervals which contained sufficient samples for statistical significance were plotted against dynamic range in decibels in figure 9. The best straight-line fit indicates good correlation of error with dynamic range, a negative error for E_{rms} and a positive error for E_{ave} and E_{log}. The error in E_{rms} at 85-db dynamic range varied from −2.7 to +2.1 with a median value of about 1 db.

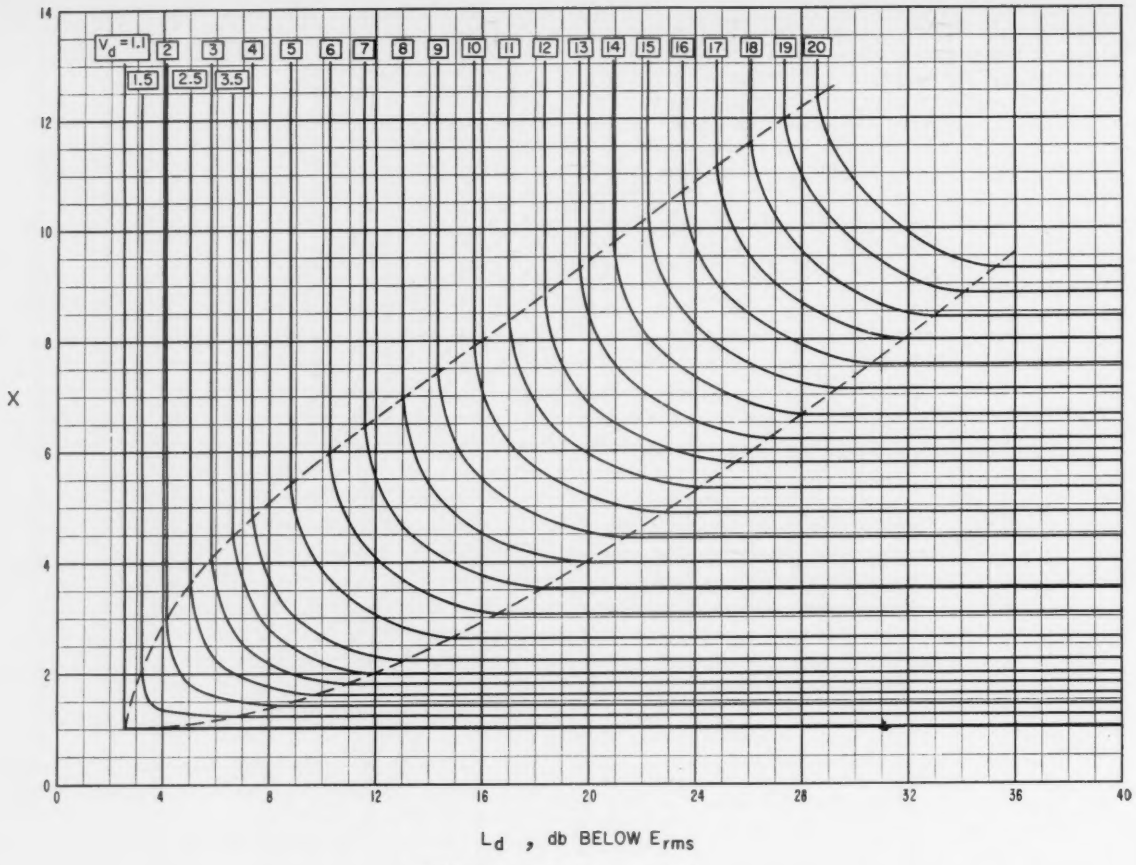


FIGURE 5. X *versus* L_d *and* V_d.

Figures 10 and 11 show the influence of errors in the moments measured by the ARN–2 on distributions determined from the moments. In these figures, the measured moments are indicated as E'_{rms} etc., and the moments obtained from the simultaneously measured distribution are indicated by the same symbols without the primes. The examples chosen contain errors typically measured by the ARN–2. In figure 10, the errors in the measured moments are small and the distribution derived from the slightly erroneous ARN–2 moments fits the measured points with negligible error. In figure 11, the dynamic range is wider with correspondingly larger errors in the ARN–2 measured moments. Thus, the departure of the distribution derived from the ARN–2 measured moments is greater. These departures, though significant are not considered excessive in view of the fact that they are of the same order of magnitude as the changes in the noise itself during an hour.

Thus, it can be concluded that this is a satisfactory method of deriving the complete amplitude-probability distribution from statistical moments, since the departures from the true distribution will be small when the moments are measured accurately.

6. Bandwidth Considerations

The foregoing method of obtaining an amplitude-probability distribution from the three measured moments results in a distribution which is only valid for the bandwidth in which the moments were measured. Fulton [8] developed a method of bandwidth conversion which was accurate for the high-amplitude, low-probability portion of the distribution, but failed to convert the lower portion of the curve as accurately since it was an empirical method based upon insufficient measurements. A study is now in progress for a more accurate method of bandwidth conversion.

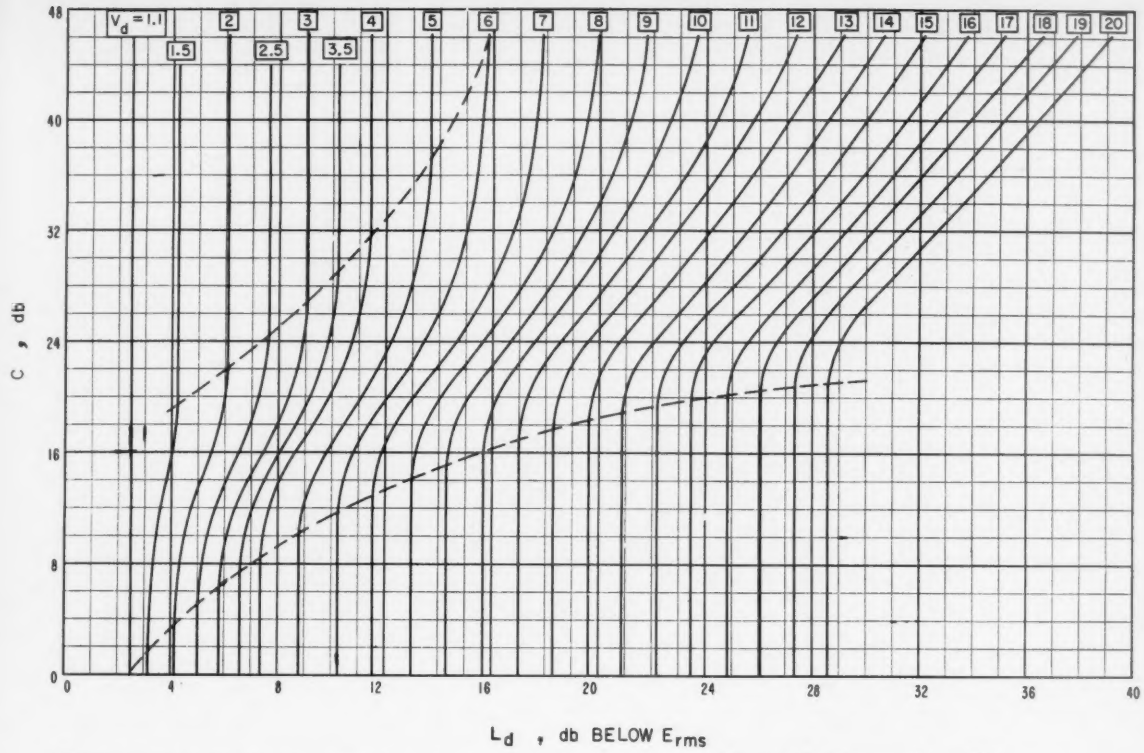


FIGURE 6. C *versus* L_d *and* V_d.

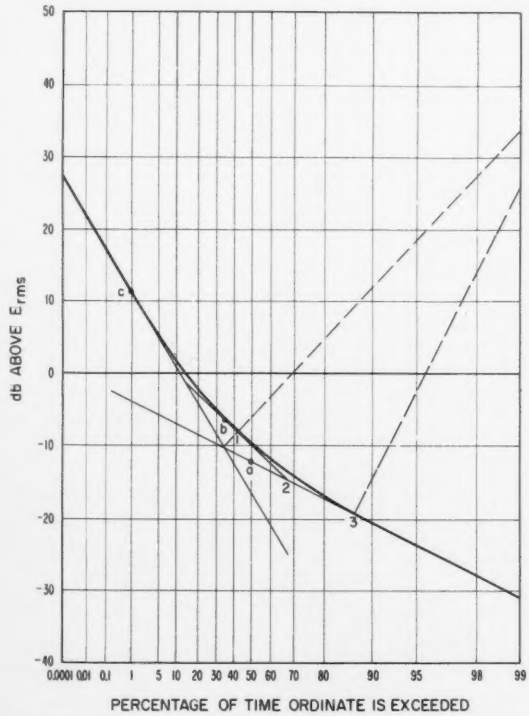


FIGURE 7. *Amplitude-probability distribution determined for* V_d=4.5 db *and* L_d=9.2 db.

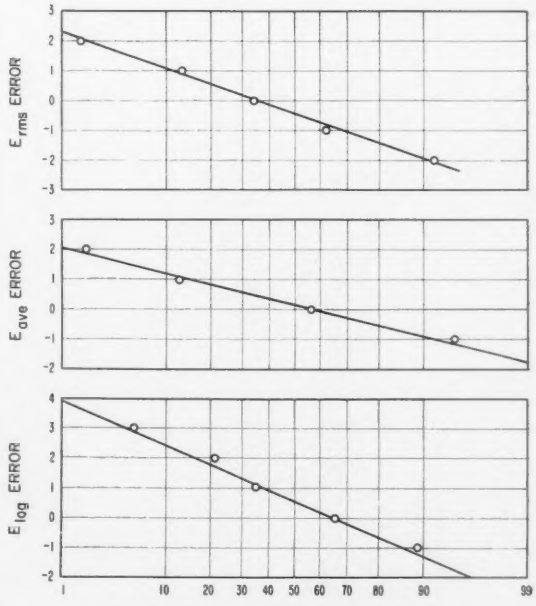


FIGURE 8. *Distribution of errors of measured moments.*

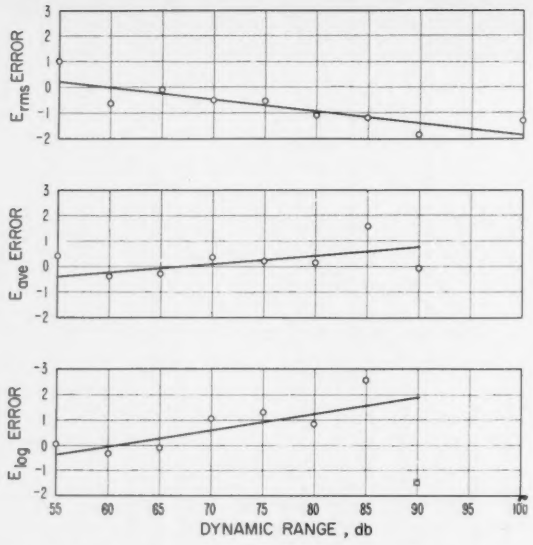


FIGURE 9. *Error of measured moments versus dynamic range.*

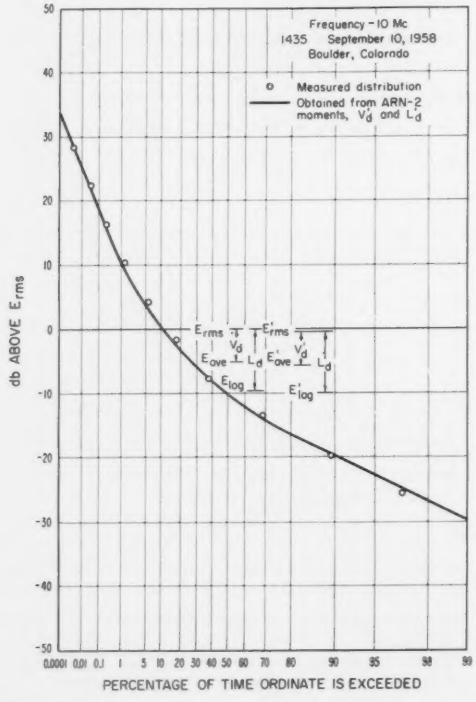

Frequency – 10 Mc
1435 September 10, 1958
Boulder, Colorado
Measured distribution
Obtained from ARN-2 moments, V_d' and L_d'
db ABOVE E_{rms}
PERCENTAGE OF TIME ORDINATE IS EXCEEDED

FIGURE 10. *Comparison of measured distribution with that obtained from moments measured by ARN–2.*

BOULDER, COLO. (Paper 64D1–37)

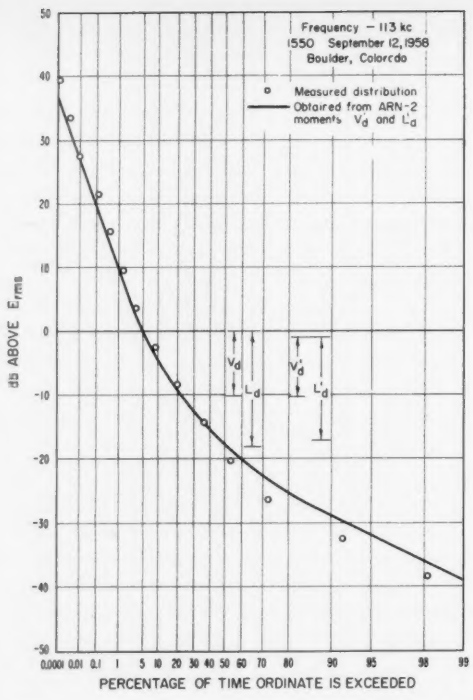

Frequency – 113 kc
1550 September 12, 1958
Boulder, Colorado
Measured distribution
Obtained from ARN-2 moments V_d and L_d
db ABOVE E_{rms}
PERCENTAGE OF TIME ORDINATE IS EXCEEDED

FIGURE 11. *Comparison of measured distribution with that obtained from moments measured by ARN–2.*

7. References

[1] International Scientific Radio Union, Recommendation I and its annex, Proceedings of the XIIth General Assembly, Commission IV, Pt. 4, **XI,** 99 (1957). (See also: H. E. Dinger, Report on URSI Commission IV—Radio noise of terrestrial origin, Proc. IRE, **46,** 1366 (1958).)

[2] W. Q. Crichlow, Report of subcommission on the question: what are the most readily measured characteristics of terrestrial radio noise from which the interference to different types of communications systems can be determined, URSI Rpt. No. 254, Proceedings of the XIIth General Assembly, Commission IV, Pt. 4, **XI,** 9 (1957).

[3] W. Q. Crichlow, Noise investigation at VLF by the National Bureau of Standards, Proc. IRE, **45,** 778 (1957).

[4] A. D. Watt and E. L. Maxwell, Measured statistical characteristics of VLF atmospheric radio noise, Proc. IRE, **45,** 55 (1957).

[5] A. W. Sullivan, The characteristics of atmospheric noise, (Atmospheric Noise Research Laboratory, Engineering and Industrial Experiment Station, University of Florida, Gainesville, Florida).

[6] H. Yuhara, T. Ishida, and M. Higashimura, Measurement of the amplitude-probability distribution of atmospheric noise, J. Radio Research Labs., **3,** 101 (1956).

[7] F. Horner, An investigation of atmospheric radio noise at very low frequencies, Proc. Inst. Elec. Engs., Pt. B, **103,** 743 (1956).

[8] F. F. Fulton, Jr., The effect of receiver bandwidth on amplitude distribution of VLF atmospheric noise, Prepublication Record, Symposium on the Propagation of VLF Radio Waves, National Bureau of Standards, Boulder, Colorado, **III,** Paper No. 37 (1957) (to be published in J. Research NBS, Sec. D).

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Measurements of Coastal Deviation of High-Frequency Radio Waves

C. W. McLeish [1]

(August 4, 1959)

The angular deviation of the phase front of a wave propagated across a fresh water shoreline has been measured over the frequency range from 3 to 20 Mc. The deviation is found to be roughly half that which theoretically would be obtained if the same sites were adjacent to infinitely conducting surfaces.

1. Introduction

Some recent measurements of coastal deviation at low frequency (lf) have been reported by Pressey, Ashwell, and Fowler [1, 2] [2]. Their results indicate the practical difficulty of interpreting measured phase front deviations. Irregular coastline and nonhomogeneous soil constants create wave interference effects which tend to mask the deviation due to the boundary.

This paper presents measurements made at high frequency (hf) (3 to 20 Mc) where the site area which affects the results can be much smaller than in the lf case. Observations were made over a number of paths at each site, thus reducing to some extent the site error.

Finally, a comparison is drawn between the experimental and the theoretical deviations which would be expected under certain ideal conditions.

2. Observations

Two sites were chosen, each with a reasonably straight shoreline along a river but with quite different soil constants. Measured values are given in table 1 for both sites and for the river water.

TABLE 1. *Site ground constants*

Site	Method	Conductivity, σ emu	Relative permittivity, K
1 arable and	Wave tilt	1×10^{-13}	9
2 limestone	Wave tilt	3×10^{-14}	25
2 rock	Surface sample	3×10^{-16}	7
River water	Sample	1×10^{-12}	81

The wave tilt measurements at site 2 are of doubtful accuracy because the path of propagation was within a half wavelength of the water at the low frequencies where a value could be obtained. However, it is likely that both conductivity and relative permittivity will be higher than that obtained in samples because of water in fissures in the rock.

A coaxial-spaced loop direction finder was used to measure the deviations of the phase front of the horizontal magnetic field. The transmitter in all cases was situated over the water. In order to eliminate systematic instrumental error, the difference between a radio bearing and the visual bearing was observed first on a path normal to the shoreline and then on a path usually at 60° to the normal. Observations were made at 1-Mc intervals in the 3- to 20-Mc band.

At site 1 readings were taken from eight positions, all approximately 40 m from the shoreline. The mean of the deviations is plotted in figure 1(a). The

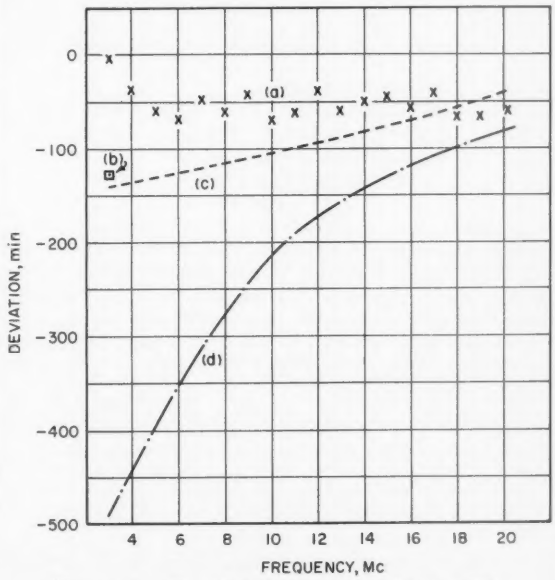


FIGURE 1. *Coastal deviations at site 1 for $\theta=60°$.*

(a) × Mean observed deviation.
(b) ⊡ Furutsu, $\sigma=1\times10^{-13}$, K=1.
(c) — — — Wait, $\sigma=1\times10^{-13}$, K=1.
(d) —·— Senior, $\sigma=0$, K=9.

rms fluctuation of individual deviations about the mean is about 43 min, indicating the effects of site error on the result.

[1] Division of Electrical and Radio Engineering, National Research Council, Ottawa, Canada.
[2] Figures in brackets indicate the literature references at the end of this paper.

At site 2, five sets of readings were obtained from positions 30 m from the shoreline and two sets at 10 m. The mean deviations are plotted in figure 2(a) and (f). The site error contribution appeared to be about the same as at site 1.

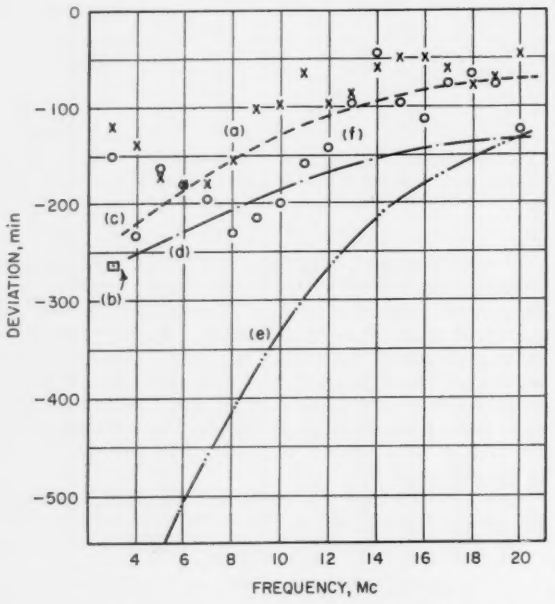


FIGURE 2. *Coastal deviations at site 2 for $\theta=60°$.*

(a) × Mean observed deviation.
(b) ⊡ Furutsu, $\sigma=3\times10^{-14}$, K=1.
(c) — — — Wait, $\sigma=3\times10^{-14}$, K=1.
(d) —·— Senior, $\sigma=0$, K=25.
(e) —··— Senior, $\sigma=0$, K=7.
(f) ⊙ Observed deviation at 10-m distance.

3. Theory

Several papers [3 to 8] deal with the refraction of a surface wave crossing the boundary between two media. The general expressions for the angular deviation of the wave front are not capable of numerical solution and in each case simplifying assumptions have to be made to arrive at values for the deviation. For instance, Furutsu [6] provides an expression for deviation when one medium is infinitely conducting, the other has finite conductivity, and the wave path numerical distance of the receiver from the boundary is small. Using parametric curves, Wait [7] has derived an expression allowing for finite conductivity on both sides of the boundary and for unrestricted distances. Senior [8] has investigated the deviation when one medium is infinitely conducting and the other a lossless dielectric. Whereas the first two cases apply chiefly to the low frequencies for groundwave propagation, the last one is applicable on higher frequencies. All the solutions require the distance from the boundary to be sufficiently large that only the radiation field is effective. Table 2 shows the expressions developed in the above papers, together with the limits imposed on each.

Apart from the dimensions defined in figure 3, where distances are in meters, the symbols used are defined as follows:

f=frequency, cps,

λ=wavelength, meters,

K=relative permittivity,

ρ_2=numerical distance along r_2,$=\frac{\pi r_2}{x}\cos b$ for vertical polarization,

$x=\frac{1.8\times10^{21}\sigma}{f}$,

$b=\text{arc tan}\,\frac{K+1}{x}$,

σ=conductivity, emu,

$V=\frac{r_2}{r_{12}}$,

$\delta V=\frac{-d\phi}{dV}$, obtained from Wait's curves of phase versus numerical distance, and

$\gamma=\frac{2\pi}{\lambda}$ and $\eta=\frac{1}{\sqrt{K}}$ in this case.

TABLE 2. *Theoretical deviations*

Author	Deviation angle	Conductivity		Distance	
	δ_2, Degrees	1st Medium	2d Medium	Numerical ρ_2	Normal $\frac{2\pi r_2\cos\theta}{\lambda}$
Furutsu	$\frac{3.73\tan\theta\sqrt{r_1/r_2}}{\sqrt{\sigma r_{12}}}$	$\frac{emu}{\infty}$	$>>\frac{fK}{1.8\times10^{21}}$	<<1	>>1
Wait	$\frac{\tan\theta\,(\delta V)(1-V)}{2\pi r_{12}}\frac{180}{\pi}$	$>>\frac{fK}{1.8\times10^{21}}$	$>>\frac{fK}{1.8\times10^{21}}$	(a)	>>1
Senior (eq 26)	$\arctan\frac{-\tan\theta}{4r_2\left(\eta\sqrt{\frac{\gamma r_2}{\pi}}+\frac{1}{2\eta}\sqrt{\frac{\pi}{\gamma r_2}}-1\right)}$	∞	$<<\frac{fK}{1.8\times10^{21}}$	------	>>1

(a) Curves are given up to $\rho_2=5.0$.

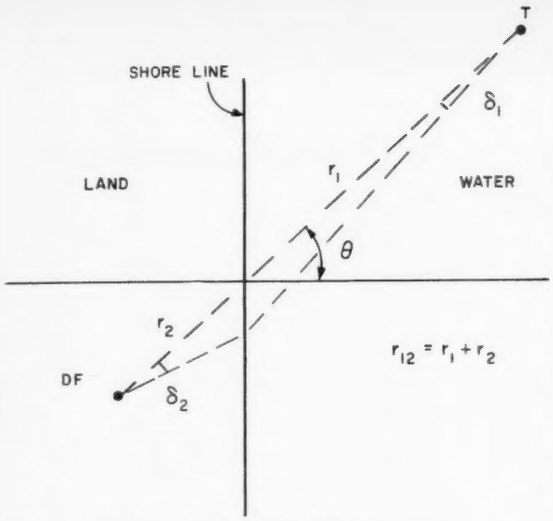


FIGURE 3. *Geometry of deviation effect.*

Senior's eq (26) refers to the vertical electric field rather than to the horizontal magnetic field deviations which are measured in the experiment.

For purposes of comparison, calculations of deviation for the three expressions have been carried out assuming the water path to be a perfect conductor in each case. Actually the water behaves as a dielectric medium above 2.2 Mc because of its low conductivity. The deviation, according to Furutsu, is plotted for both sites in figures 1(b) and 2(b) for 3 Mc only, for at higher frequencies the numerical path ρ_2 is no longer small. In the same figures the curves (c) and (d) are for Wait's and Senior's expressions, respectively, using the indicated values of conductivity and permittivity. Curve (e) in figure 2 is for the measured sample value of relative permittivity.

4. Discussion of Results

At site 1 there is sufficiently high conductivity that the propagation constant does not become dependent on permittivity until the frequency is above 20 Mc. Therefore, the observed deviations should approach those given by Wait, except near the low end of the band where the water propagation constant is changing rapidly to that of a lossy medium having the same conductivity as the site. This may be the reason for the observed reduction of deviation at 3 Mc.

At site 2 propagation over the whole frequency range is affected chiefly by its permittivity. Therefore, one would expect deviations approaching those given by Senior. The observed deviations have a similar trend above 6 Mc but are considerably smaller than expected for the ideal case. The difference may be partly because the measurements refer to the horizontal magnetic field while the theoretical curve refers to the vertical electric field, and partly because the water permittivity is finite. The drop of deviation below 6 Mc may again be attributed to the change of propagation over the water at low frequencies.

The same general trend is also apparent in figure 2(f) and the deviations are somewhat larger. The distance from the shoreline in this case is, of course, too small to permit comparison with any of the theoretical expressions.

5. Conclusions

In the siting of direction finding stations and similar navigational aids near bodies of water, deviations of the wave front may introduce error in observations. All three theoretical expressions show that the angular deviation of the wave is approximately proportional to tan θ and inversely proportional to r_2. Therefore, at large values of θ, where the error may be considered significant, it is approximately inversely proportional to $r_2 \cos \theta$, the normal distance of the receiver from the boundary. By extrapolation from the observed results one might expect deviation errors on a high conductivity site adjacent to fresh water to be less than ½ deg if the distance from the shoreline were at least 80 m. Similarly on a site of very low conductivity, where the observed errors had a definite frequency dependence, the required distance for the same error limit is about 4 wavelengths. By comparison, sites beside the sea would, according to the theory, have to be about twice as far from the shoreline.

The author thanks J. McDougall and J. Lee for carrying out the experimental work.

6. References

[1] B. G. Pressey, G. E. Ashwell, C. S. Fowler, Change of phase with distance of a low frequency ground wave propagated across a coast line, Proc. Inst. Elec. Engrs., Paper 2082 R **103,** Pt. B, 527 (July 1956).

[2] B. G. Pressey, G. E. Ashwell, The deviation of low frequency ground waves at a coastline, Proc. Inst. Elec. Engrs., Paper 2083 R **103,** Pt. B, 535 (July 1956).

[3] P. C. Clemmow, Radio propagation over a flat earth across a boundary separating two different media, Phil. Trans. Roy. Soc. London, Ser. A **246,** 1 (June 1953).

[4] G. Grunberg, Suggestion for a theory of coastline refraction, Phys. Rev. **63,** 185 (1943).

[5] E. Feinberg, On the propagation of radio waves along an imperfect surface, J. Phys. (U.S.S.R.) **10,** 410 (1946).

[6] K. Furutsu, Propagation of electromagnetic waves over a flat earth across a boundary separating different media and coastal refraction. J. Radio Research Labs. (Japan) **2,** 1 (June 1955).

[7] J. R. Wait, Mixed path ground wave propagation: 1. Short distances, J. Research NBS **57,** 1 (July 1956) RP2687; and J. R. Wait, J. Householder, 2. Larger distances, J. Research NBS **59,** 19 (July 1957) RP2770.

[8] T. B. A. Senior, Radio propagation over a discontinuity in the earth's electrical properties—II. Coastal refraction, Proc. Inst. Elec. Engrs., Paper 201 R **104,** Pt. C, 139 (March 1957).

BOULDER COLO. (Paper 64D1–38)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

An Exact Earth-Flattening Procedure in Propagation Around a Sphere[1]

Benjamin Y.-C. Koo[2] and Martin Katzin[2]

(August 17, 1959)

By a refinement of the procedure used in the usual earth-flattening approximation, the problem of propagation around a spherical earth is reduced to an exact equation of the same form. Thereby the earth-flattening procedure becomes applicable to arbitrarily large heights and distances. It is also found that existing solutions of the approximate equations can be re-evaluated to yield the exact solutions for slightly different refractive index distributions.

1. Introduction

The problem of propagation of radio waves around a spherical earth has been attacked by many writers over a period of more than a half a century. For the idealized case of homogeneous atmosphere and earth, the presently accepted solution is due to van der Pol and Bremmer [1]. This makes use of a modified form of the Watson transformation to convert from a very slowly converging series of spherical harmonics to a more rapidly converging series of residue waves. Physically, the transformation is from waves progressing radially out from the earth to waves which progress along the earth's surface.

In order to reduce the mathematical complexities accompanying the van der Pol—Bremmer treatment, M. H. L. Pryce [2, 3] introduced the so-called *earth-flattening approximation*, in which plausible approximations are introduced to simplify the original differential equations. This procedure reduces the problem of propagation over spherical earth to propagation over a plane earth with an atmosphere having a *modified refractive index*. In this way the van der Pol—Bremmer type of result for the homogeneous atmosphere can be obtained in a much more direct manner. Furthermore, the method can be applied to inhomogeneous atmospheres in which the refractive index varies only with height (so-called *horizontal stratification*).

Pekeris [4] investigated the accuracy of the earth-flattening approximation in the case of the *standard atmosphere* (in which the modified refractive index increases linearly with height). He found that the accuracy decreased with increasing distance along the earth's surface and also with increasing height above the surface. If r is the great-circle distance and a the earth's radius, the error with distance is proportional to $(r/a)^2$ independently of the frequency, amounting to about 8 percent at $r/a=1$. The error with height, however, is proportional to frequency, and to height as $h^{5/2}$. The height error becomes serious in the microwave region at heights in the order of a few thousand feet.

In this paper we improve and extend the Pryce and Pekeris treatments to obtain exact differential equations for the spherical geometry in terms of equations of plane earth type. The solutions of these hold for arbitrary heights and distances. Comparison with the previous earth-flattening approximation leads to an interpretation of the errors therein.

2. Formulation of the Problem

We consider propagation from a source above a spherical earth of radius a with an atmosphere whose refractive index n is a function only of the radial distance (i.e., horizontally-stratified atmosphere). The treatment is simplified without any essential restriction by assuming that radiation from the source is independent of azimuth (we shall indicate later the extension to remove this restriction). Then it can be shown [5] that the electric and magnetic fields at any point may be derived from a radial Hertz vector ψ, which satisfies the reduced wave equation

$$\nabla^2\psi+k_0^2n^2\psi=0, \qquad (2.1)$$

where $k_0=2\pi/\lambda=2\pi f/c$. By adopting a spherical coordinate system (R,θ,ϕ) with origin at the center of the earth and with the source on the polar axis, (2.1) (with the assumed azimuth independence of ψ) reduces to

$$\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial\psi}{\partial R}\right)+\frac{1}{R^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+k_0^2n^2\psi=0. \qquad (2.2)$$

This equation can be separated in the usual way by letting

$$\psi=T(\theta)U(R),$$

[1] The research reported in this paper was supported in part by the Missile and Space Vehicles Department of the General Electric Co.
[2] Electromagnetic Research Corp., Washington, D.C.

whereby there result the two ordinary differential equations

$$\frac{d^2U}{dR^2}+\frac{2}{R}\frac{dU}{dR}+k_0^2\left[n^2-\left(\frac{aS}{R}\right)^2\right]U=0, \tag{2.3}$$

$$\frac{d^2T}{d\theta^2}+\cot\theta\frac{dT}{d\theta}+(as)^2T=0, \tag{2.4}$$

in which $(as)^2$ is the separation constant and $S=s/k_0$.

We now introduce the transformations [3]

$$\eta=a\log(R/a), \tag{2.5}$$

$$r=a\theta. \tag{2.6}$$

Equations (2.3) and (2.4) become, respectively,

$$\frac{d^2U}{d\eta^2}+\frac{1}{a}\frac{dU}{d\eta}+k_0^2(n^2e^{2\eta/a}-S^2)U=0, \tag{2.7}$$

$$\frac{d^2T}{dr^2}+\frac{1}{a}\cot(r/a)\frac{dT}{dr}+s^2T=0. \tag{2.8}$$

S^2 is the eigenvalue of the differential eq (2.7), which eventually is determined by the boundary conditions.

The problem of solving (2.1) thus is equivalent to solving (2.7) and (2.8).

3. The Height-Gain Function U

We consider first (2.7) for the height-gain function U. By introducing a new variable V, given by

$$U=(a/R)^{\frac{1}{2}}V=Ve^{-\eta/(2a)}, \tag{3.1}$$

(2.7) becomes

$$\frac{d^2V}{d\eta^2}+k_0^2(N^2-\sigma^2)V=0, \tag{3.2}$$

in which we introduce as our definition of the modified refractive index

$$N=ne^{\eta/a}=nR/a,$$

and define

$$\sigma^2=S^2+1/(2k_0a)^2.$$

(3.2) is exactly the same *in form* as the plane-earth equation, which is

$$\frac{d^2U_0}{dh^2}+k_0^2(\overline{N}^2-S^2)U_0=0, \tag{3.3}$$

where $h=R-a$ is the actual height above the earth's surface, and

$$\overline{N}^2=n^2(1+h/a)^2\approx n^2(1+2h/a).$$

[3] These are the transformations used by Pryce [3], who credited Pekeris for the range transformation and E. T. Copson for the height transformation.

The method of the usual earth-flattening approximation is to approximate U by U_0, which is the solution of (3.3). But U can be obtained exactly by means of (3.1), in which V is the solution of (3.2). This fact will permit us to interpret properly the significance of the earth-flattening approximation with respect to the height dependence.

4. Interpretation of the Earth-Flattening Approximation

Equations (3.2) and (3.3) are identical only in form. However, the height variable η differs from h, and, for a given atmosphere, N as a function of η will not have the same analytical form as $\overline{N}$ as a function of h. In addition, the eigenvalue σ^2 differs from S^2 by the small real quantity $1/(2k_0a)^2$. Thus it follows that a given solution U_0 of the approximate earth-flattened eq (3.3) corresponds (but is not equal) to the exact solution of a slightly different problem. The approximate solution differs from the corresponding exact solution in three respects, namely: (a) By the factor $(a/R)^{1/2}$ in (3.1) (this factor usually is nearly unity); (b) the modified index profile is different because of the difference in height scale; [4] (c) the eigenvalue is shifted slightly. Both (b) and (c) lead to differences between U_0 and U which involve the frequency. (b) probably is the more important effect, so that *the error of the earth-flattening approximation as a function of height probably is due primarily to the choice of height scale.*

5. The Standard Atmosphere

The standard atmosphere is now defined by

$$\overline{N}^2(h)=1+qh,$$

$$q=3/(2a).$$

In this case, (3.3) becomes Stokes' equation, whose solutions are Airy functions, or the modified Hankel functions of order one-third [6]. In view of the fact that an exact solution is available in terms of our function U, it appears that it may be desirable to define the standard atmosphere by the relation

$$N^2(\eta)=1+q\eta.$$

In this case (3.2) likewise becomes Stokes' equation.

6. The Distance Function T

The distance dependence is given by the function T, which satisfies the Legendre differential eq (2.8). In the earth-flattening approximation, T is approximated by a function T_0 which satisfies the differential equation

$$\frac{d^2T_0}{dr^2}+\frac{1}{r}\frac{dT_0}{dr}+s^2T_0=0, \tag{6.1}$$

[4] This shall be discussed in greater detail for specific cases in a later paper.

which is the Bessel equation of order zero. Equation (6.1) gives the distance dependence for plane earth.

In investigating the accuracy of this approximation, Pekeris [4] assumed that T could be written in the form

$$T(r)=T_0(r)+(as)^{-2}T_2(r)+(as)^{-4}T_4(r)+\ldots. \quad (6.2)$$

After substituting (6.2) into (2.8), expanding $\cot(r/a)$ in a power series in r/a and equating the coefficients of like powers of $(as)^{-2}$, he obtained a system of simultaneous differential equations, which is written in terms of the operator

$$L(T)=\left(\frac{d^2}{dx^2}+\frac{1}{x}\frac{d}{dx}+1\right)T,$$

as follows:

$$L(T_0)=0, \quad (6.3a)$$

$$L(T_2)=\frac{x}{3}\frac{dT_0}{dx}, \quad (6.3b)$$

$$L(T_4)=\frac{x^3}{45}\frac{dT_0}{dx}+\frac{x}{3}\frac{dT_2}{dx}, \quad (6.3c)$$

$$L(T_6)=\frac{2x^5}{945}\frac{dT_0}{dx}+\frac{x^3}{45}\frac{dT_2}{dx}+\frac{x}{3}\frac{dT_4}{dx}, \quad (6.3d)$$

where $x=sr$.

Pekeris wrote the solutions of (6.3a) and (6.3b) in terms of the Hankel functions (corresponding to an implicit time function $e^{i\omega t}$)

$$T_0=H_0^{(2)}(sr),$$

$$T_2=\frac{(sr)^2}{12}H_2^{(2)}(sr),$$

and used

$$(as)^{-2}T_2(r)=-(1/12)(r/a)^2H_2^{(2)}(sr)$$

as the term to estimate the error in using T_0 as the solution of (2.8).

We now proceed to extend Pekeris' treatment to solve for $T_{2n}(r)$ for arbitrary n. The solution will be obtained in terms of cylinder functions of correspondingly higher order. Hence it is possible to obtain $T(r)$ in the form (6.2) to any desired degree of accuracy for any distance r.

Let $Z_n(x)$ be any cylinder function of order n. Define

$$C_n(x)\equiv x^nZ_n(x)$$

and

$$L(T)\equiv T''+x^{-1}T'+T,$$

where primes indicate derivatives with respect to x. Then C_n satisfies the differential equation.

$$L(C_n)=(2n/x)C_n'. \quad (6.4)$$

Now a solution of (6.3a) is

$$T_0=C_0=Z_0, \quad (6.5a)$$

and since

$$C_0'=-x^{-2}C_2',$$

(6.3b) becomes

$$L(T_2)=-(1/3)x^{-1}C_2'=-(1/12)4x^{-1}C_2'.$$

Hence from (6.4) it follows that

$$T_2=-(1/12)C_2=-(1/12)(sr)^2Z_2(sr). \quad (6.5b)$$

Similar procedures may be applied to (6.3c), (6.3d), . . . by applying the following reduction formula to reduce the terms on the right-hand side to a sum of terms of the form $2nx^{-1}C_n'$:

$$x^{2k}C_m'=\sum_{\alpha=0}^{k}C_{\alpha,k,m}C_{m+k+\alpha}', \qquad k=1,2,\ldots, \quad (6.6)$$

where

$$\left.\begin{aligned} &C_{0,k,0}=0,\\ &C_{\alpha,k,m}=(-1)\frac{\alpha k!(m+k-1)!\,2^{k-\alpha}}{\alpha!(k-\alpha)!(m+\alpha-1)!}, \qquad m+\alpha\neq 0. \end{aligned}\right\} \quad (6.7)$$

Equation (6.6) may be derived from the recurrence formulas for the cylinder functions. The solutions for T_4 and T_6 are given below:

$$T_4=-\frac{1}{30}C_3+\frac{1}{160}C_4, \quad (6.5c)$$

$$T_6=-\frac{17}{840}C_4+\frac{19}{2520}C_5-\frac{61}{120{,}960}C_6. \quad (6.5d)$$

In most cases of propagation around the earth, one or two terms of (6.2) give sufficient accuracy.

Finally, the treatment may be extended in a similar way to the case of $\partial\psi/\partial\phi\neq 0$. Equation (2.4) then is replaced by the associated Legendre equation, for which we have obtained a solution in terms of generalized C functions

$$C_{n,m}(x)=x^nZ_{m+n}(x).$$

The details of this development will be reported elsewhere.

7. Conclusions

By a refinement of the procedure used in the usual earth-flattening approximation, we have shown that the problem of propagation around a spherical earth can be reduced to an exact equation of the same form. Thereby the earth-flattening procedure becomes applicable to arbitrarily large heights and distances. It is also found that existing solutions of the approximate equations can be re-evaluated to yield the exact solutions for slightly different refractive index distributions.

8. References

[1] B. van der Pol and H. Bremmer, The diffraction of electromagnetic waves from an electrical point source round a finitely conducting sphere, with applications to radiotelegraphy and the theory of the rainbow, Phil. Mag. **24,** 141, 825 (Suppl.) (1937); The propagation of radio waves over a finitely conducting spherical earth, Phil. Mag. **25,** 817 (Suppl.) (1938); Further note on the propagation of radio waves over a finitely conducting spherical earth, Phil. Mag. **27,** 261 (1939).

[2] M. H. L. Pryce (unpublished).

[3] M. H. L. Pryce, The diffraction of radio waves by the curvature of the earth, Advances in Phys. **2,** 67 (1953).

[4] C. L. Pekeris, Accuracy of the earth-flattening approximation in the theory of microwave propagation, Phys. Rev. **70,** 518 (1946).

[5] H. Bremmer, Terrestrial radio waves, p. 139 (Elsevier Publishing Co., Inc., Amsterdam, Netherlands, 1949).

[6] Computation Laboratory, Harvard University, Tables of the modified Hankel functions of order one-third and of their derivatives (Harvard University Press, Cambridge, Massachusetts, 1945).

BOULDER, COLO.

(Paper 64D1–39)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Limit of Spatial Resolution of Refractometer Cavities

William J. Hartman

(July 9, 1959)

Filter factors that determine an upper limit for the wave numbers for which refractometer measurements can be used to calculate the spectrum of refractivity are derived in this paper based on the assumption that refractometers measure a weighted average of the refractive index in a volume of air surrounding the center of the refractometer cavity. Two models are used assuming the weighting function has spherical symmetry around the center and one model is used assuming the function has cylindrical symmetry. All models result in a simple mathematical form which should be easy to use in further theoretical developments.

1. Introduction

The wave number power spectrum of refractivity is determined for large wave numbers k by the correlation of the refractive index variations at points which are separated by small distances [1].[1] Thus, refractometer measurements discriminate in two ways against the spectrum for large wave numbers. First, the refractometer does not measure the variation of the refractive index at a point, but rather, a weighted average of the variations in the volume of air in the vicinity of the center of the refractometer cavity, and second, the centers of two cavities may not be placed arbitrarily close because of obvious physical limitations. The first of these effects is evaluated in this paper assuming several different weighting functions and neglecting further practical limitations such as finite response time and noise characteristics of the instrument. The result shows that the measured spectrum is a filter factor times the theoretical spectrum. With appropriate approximations, the effect is also evaluated for the measured frequency spectrum of refractivity.

The exact weighting function is determined by, among other things, the size and shape of the cavity and the size and shape of the openings which allow the passage of air through the cavity. The commonly used microwave refractometer cavity [2] consists of a right circular cylinder with the ends partially opened to permit the flow of air and operated in the TE_{011} mode at about 9 kMc. Considering the weighting function for a closed cylindrical cavity operated in the TE_{011} mode, it appears that the weighting function for an open cavity with the same dimensions, and operated in the same mode, decreases approximately as a partial sine wave across a section of the cavity perpendicular to the axis and gaussianly along the axis from the center [3]. For mathematical simplicity, however, the most desirable weighting function is one that has spherical symmetry about the center of the cavity, as will be seen later in the analysis in this paper. (This would be a very good approximation, for example, for a refractometer using a cavity consisting of two concentric spheres, the outer sphere being a grid to allow for the passage of air.) Two of the weighting functions used in this paper have spherical symmetry, and a third function is used which has cylindrical symmetry. The results for the three different forms differ significantly only in the region where data become unreliable because of other limitations.

2. Theoretical Development

It may be assumed that the refractometer measures a weighted spatial average of the variation $\Delta n = n - n_0$ where n_0 denotes an average value of the refractive index n. Let the

[1] Figures in brackets indicate the literature references at the end of this paper.

524214—59——5

weighted average of Δn at any time t be given by

$$\overline{\Delta n}(\vec{R},t)=\frac{\int_v d^3r\Delta n(\vec{r}',t)f(\vec{r})}{\int_v d^3rf(\vec{r})}=\frac{\int_v d^3r\Delta n(\vec{R}+\vec{r},t)f(\vec{r})}{\int_v d^3rf(\vec{r})} \tag{1}$$

where v is a volume centered at $\vec{R}$, the center of the refractometer cavity, and $f(\vec{r})$ is the weighting function for the cavity. Figure 1 shows two spherical cavities and the vectors associated with these.

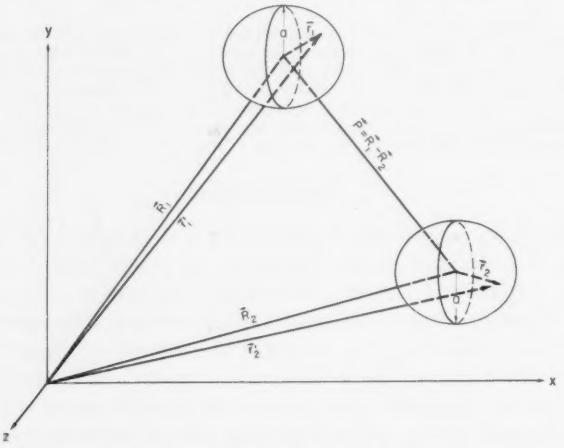


FIGURE 1. *Geometry for two spherical refractometer cavities.*

Using two refractometers with centers at $\vec{R}_1$ and $\vec{R}_2$, it follows that

$$<\overline{\Delta n}(\vec{R}_1,t)\overline{\Delta n}(\vec{R}_2,t)>=\frac{\int_{v_1} d^3r_1\int_{v_2} d^3r_2<\Delta n(\vec{r}_1',t)\Delta n(\vec{r}_2',t)>f(\vec{r}_1)f(\vec{r}_2)}{\int_{v_1} d^3r_1\int_{v_2} d^3r_2f(\vec{r}_1)f(\vec{r}_2)} \tag{2}$$

where $<\ >$ denotes a time average or ensemble average.

Following Wheelon [4][2], one obtains

$$<\Delta n(\vec{r}_1',t)\Delta n(\vec{r}_2',t)>=<\Delta n^2>C(\vec{r}_1'-\vec{r}_2')=\frac{1}{8\pi^3}\int_v d^3kS(\vec{k})\ \exp\ [-i\vec{k}\cdot(r_1'-r_2')] \tag{3}$$

where $<\Delta n^2>$ is the variance of the fluctuations of the refractive index. $C(\vec{r}_1'-\vec{r}_2')$ is the correlation of the fluctuations at two points determined by the vectors $\vec{r}_1'$ and $\vec{r}_2'$, and $S(\vec{k})$ is the wave number power spectrum of refractivity.

Similarly, one may define

$$<\overline{\Delta n}(\vec{R}_1,t)\overline{\Delta n}(\vec{R}_2,t)>=<\Delta n^2>C_w(\vec{P}) \tag{4}$$

where $C_w(\vec{P})$ is the correlation function between the centers of the cavities of two refractometers, and the vector $\vec{P}$ is the vector between the two centers, $\vec{P}=\vec{R}_1-\vec{R}_2$.

[2] The magnitude of the $S(k)$ defined by Wheelon is about one-fourth as large as the one used here since he used $\Delta\epsilon\cong 2n\Delta n$ instead of Δn.

Combining (2), (3), and (4) one finds that

$$<\overline{\Delta n}(\vec{R}_1, t)\overline{\Delta n}(\vec{R}_2, t)> = <\Delta n^2> C'_w(\vec{P}) = \frac{\int_{v_1} d^3r_1 \int_{v_2} d^3r_2 <\Delta n^2> C(\vec{r}_1' - \vec{r}_2') f(\vec{r}_1) f(\vec{r}_2)}{\int_{v_1} d^3r_1 \int_{v_2} d^3r_2 f(\vec{r}_1) f(\vec{r}_2)}$$

$$= \frac{\int_{v_1} d^3r_1 \int_{v_2} d^3r_2 f(\vec{r}_1) f(\vec{r}_2) \left[\frac{1}{8\pi^3} \int_v d^3k \exp\left[-i\vec{k}\cdot(\vec{r}_1' - \vec{r}_2')\right] S(\vec{k}) \right]}{\int_{v_1} d^3r_1 \int_{v_2} d^3r_2 f(\vec{r}_1) f(\vec{r}_2)}$$

$$= \frac{1}{8\pi^3} \int_v d^3k \exp\left[-i\vec{k}\cdot\vec{P}\right] S(\vec{k}) \left\{ \frac{\int_{v_1} d^3r_1 \int_{v_2} d^3r_2 f(\vec{r}_1) f(\vec{r}_2) \exp\left[-i\vec{k}\cdot(\vec{r}_1 - \vec{r}_2)\right]}{\int_{v_1} d^3r_1 \int_{v_2} d^3r_2 f(\vec{r}_1) f(\vec{r}_2)} \right\}. \tag{5}$$

The factor in the braces in (5) is called the filter factor E and this determines the upper limit of k for which the refractometer may be used to measure $S(\vec{k})$. It should be noted that in the above formulation, the exact form of the structure of the atmospheric turbulence, whether isotropic or anisotropic, affects the weighting function but will have no other effect on E. In practice, P must be greater than the size of the cavity, but for theoretical purposes, one may assume that P can take on any value greater than or equal to zero. Then, taking the inverse Fourier transform of (5), one finds

$$ES(\vec{k}) = \int_v d^3P \exp\left[i\vec{k}\cdot\vec{P}\right] <\Delta n^2> C_w(\vec{P}) \tag{6}$$

which is the quantity obtained from measurements.

In the following we will evaluate E assuming the following forms for $f(\vec{r})$:

$$1°\ f(\vec{r}) = \begin{cases} 1 \text{ for } |\vec{r}| \le a, \\ 0 \text{ for } |\vec{r}| > a \end{cases}$$

$$2°\ f(\vec{r}) = \exp\{-r^2/a^2\},$$

and

$$3°\ f(\vec{r}) = \begin{cases} \exp\{-r_x^2/b^2\} & \text{for } \sqrt{r_y^2 + r_z^2} \le a \\ 0 & \text{for } \sqrt{r_y^2 + r_z^2} > a, \end{cases}$$

where r_x, r_y, and r_z are the components of $\vec{r}$ in the x, y, and z directions respectively. The first form, 1°, gives an upper bound for any closed spherical refractometer cavity of radius a, and the second form, 2°, gives an upper bound for such a cavity operated in the TE_{011} mode. In each of the three forms it is possible (but not necessary) to identify the parameter a (or b) with the dimension of the cavity.

For the first case, 1°, if one has two refractometers which measure a small spherical volume of air of radius a and b, respectively

$$E = \left\{ \frac{1}{\left(\frac{4}{3}\pi a^3\right)} \int_0^\pi \sin\lambda_1 d\lambda_1 \int_0^{2\pi} d\theta_1 \int_0^a r_1^2 dr_1 \exp\left[-ikr_1 \cos\lambda_1\right] \right\}$$

$$\times \left\{ \frac{1}{\left(\frac{4}{3}\pi b^3\right)} \int_0^\pi \sin\lambda_2 d\lambda_2 \int_0^{2\pi} d\theta_2 \int_0^b r_2^2 dr_2 \exp\left[ikr_2 \cos\lambda_2\right] \right\}. \tag{7}$$

The quantities in the braces in (7) are the same, except that the parameter (a) is changed and one is the complex conjugate of the other. Thus we may write

$$E=g(k,a)g^*(k,b) \tag{8}$$

where $g^*(k,b)$ is the complex conjugate of $g(k,b)$. One then obtains by integrating

$$g(k,a)=\frac{3}{(ka)^3}\left[\sin ka-ka\cos ka\right] \tag{9}$$

and thus,

$$E=\frac{9}{(ka)^3(kb)^3}\left[\sin ka-ka\cos ka\right]\left[\sin kb-kb\cos kb\right]. \tag{10}$$

For the second case, 2°, one finds that

$$E=\left\{\frac{1}{a^3\pi^{3/2}}\int_0^{\pi}\sin\lambda_1 d\lambda_1\int_0^{2\pi}d\theta_1\int_0^{\infty}r_1^2dr_1\exp\left[-ikr_1\cos\lambda_1\right]\exp\left\{-r_1^2/a^2\right\}\right\}$$
$$\times\left\{\frac{1}{b^3\pi^{3/2}}\int_0^{\pi}\sin\lambda_2 d\lambda_2\int_0^{2\pi}d\theta_2\int_0^{\infty}r_2^2dr_2\exp\left[ikr_2\cos\lambda_2\right]\exp\left\{-r_2^2/b^2\right\}\right\}. \tag{11}$$

Again, one may set the two factors of the right hand side of (11) equal respectively to $g(k,a)$ and $g^*(k,b)$. After performing the λ and θ integrations, one obtains

$$g(k,a)=\frac{4}{a^3\sqrt{\pi}}\int_0^{\infty}\frac{\sin kr}{kr}r^2\exp\left\{-r^2/a^2\right\}dr \tag{12}$$

which yields, using relation No. 19, page 73, of reference[5],

$$g(k,a)=\frac{4}{a^3\sqrt{\pi}}\frac{a^3\sqrt{\pi}}{4}\frac{k}{k}\exp\left\{-\frac{a^2k^2}{4}\right\}=\exp\left\{-\frac{a^2k^2}{4}\right\}. \tag{13}$$

Thus

$$E=\exp\left\{-k^2(a^2+b^2)/4\right\}.$$

Figure 2 shows the filter factors for the sharp spherical cutoff $E_s(ka)$ and for the gaussian cutoff $E_g(ka)$ versus ka, assuming that the refractometers used to determine E_s (or E_g) are the same.

For the case 3°, E is given by

$$E=g(a,\,b)g^*(c,\,d) \tag{14}$$

where

$$g(a,\,b)=\frac{\displaystyle\int_{-\infty}^{\infty}dr_x\int_{-a}^{a}dr_y\int_{-L}^{L}dr_z\,e^{-\vec{ik}\cdot\vec{r}}\exp\left\{-r_x^2/b^2\right\}}{\displaystyle\int_{-\infty}^{\infty}dr_x\int_{-a}^{a}dr_y\int_{-L}^{L}dr_z\exp\left\{-r_x^2/b^2\right\}} \tag{15}$$

with $L=\sqrt{a^2-r_y^2}$ and $g^*(c,\,d)$ is the complex conjugate of $g(c,\,d)$. After integrating the denominator of (15) one obtains

$$g(a,\,b)=\frac{1}{\pi^{\frac{3}{2}}a^2b}\int_{-\infty}^{\infty}dr_x\int_{-a}^{a}dr_y\int_{-L}^{L}dr_z\exp\left[-i(r_xk_x+r_yk_y+r_zk_z)\right]\exp\left\{-r_x^2/b^2\right\}. \tag{16}$$

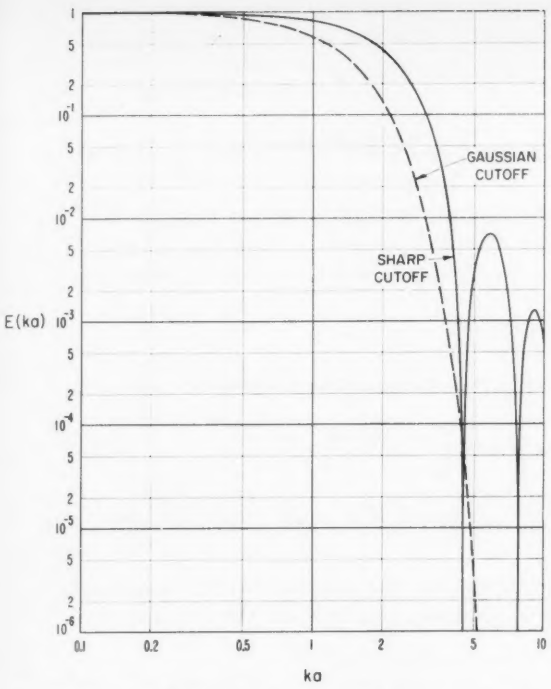


FIGURE 2. *Filter factors for gaussian cutoff and sharp cutoff.*

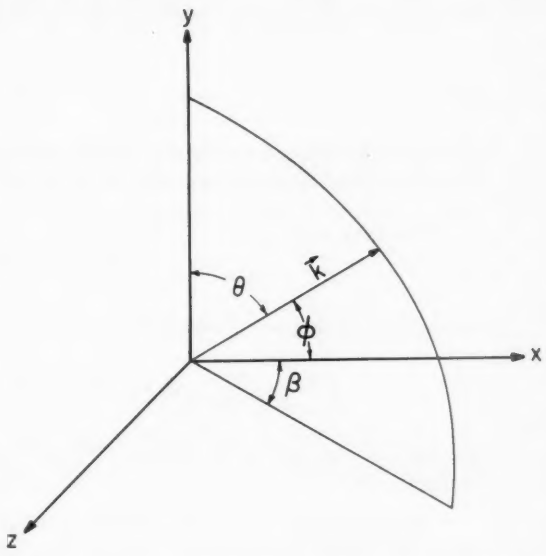


FIGURE 3. *Geometry for locating the propagation vector* $\vec{k}$ *in relation to a cylindrical cavity with its axis parallel to the x-axis.*

The r_x integration may be done using relation No. 23, page 121 in reference [5] and one obtains

$$g(a,b)=\frac{\exp\{-k_x^2b^2/4\}}{\pi a^2}\int_{-a}^{a}dr_y\int_{-L}^{L}dr_z\exp\{-i(r_yk_y+r_zk_z)\}. \tag{17}$$

Changing to polar coordinates, (17) becomes

$$g(a,b)=\frac{\exp\{-k_x^2b^2/4\}}{\pi a^2}\int_0^a ydy\int_0^{2\pi}d\theta\exp[-iy\sqrt{k_y^2+k_z^2}\sin\theta], \tag{18}$$

which may be integrated to yield

$$g(a,b)=\frac{2}{a\sqrt{k_y^2+k_z^2}}J_1(a\sqrt{k_y^2+k_z^2})\exp\left[\frac{-k_x^2b^2}{4}\right], \tag{19}$$

where $J_1(Z)$ is the Bessel function of order 1. Referring to figure 3, one finds that

$$\sqrt{k_y^2+k_z^2}=k\sqrt{1-\sin^2\theta\cos^2\beta} \tag{20a}$$

and

$$k_x^2=k^2\sin^2\theta\cos^2\beta, \tag{20b}$$

where θ is the angle which $\vec{k}$ makes with the vertical, and β is the angle in the XZ plane with the x-axis assuming that the cavities are both alined along the x-axis. Combining (14), (19),

(20a), and (20b) it follows that

$$E_c=\frac{4}{ack^2(1-\sin^2\theta\cos^2\beta)}\,J_1(ak\sqrt{1-\sin^2\theta\cos^2\beta})\,J_1\,(ck\sqrt{1-\sin^2\theta\cos^2\beta})$$

$$\times\exp\left\{-\frac{(\sin^2\theta\cos^2\beta)k^2(b^2+d^2)}{4}\right\}. \quad (21a)$$

E_c is plotted versus ka in figure 4, assuming two identical cavities, and also that a equals b in (21a), for the values β equal zero and several values of θ.

One may also write

$$\sqrt{1-\sin^2\theta\cos^2\beta}=\cos\phi,$$

where ϕ is the angle $\vec{k}$ forms with the x-axis. With this notation, (21a) becomes

$$E_c=\frac{4}{ack^2\cos^2\phi}\,J_1\,(ak\cos\phi)\,J_1\,(ck\cos\phi)\exp\left\{-\frac{k^2(b^2+d^2)\sin^2\phi}{4}\right\}. \quad (21b)$$

In figure 4, the curves for θ equal to 0°, 45°, and 90° correspond to ϕ equal to 90°, 45° and 0°, respectively.

If one assumes isotropic turbulence and then carries out the angular integrations in (5) using the filter factor (21a), a new filter factor is obtained which, as in the two previous forms, is dependent only on the magnitude but not on the direction of $\vec{k}$.

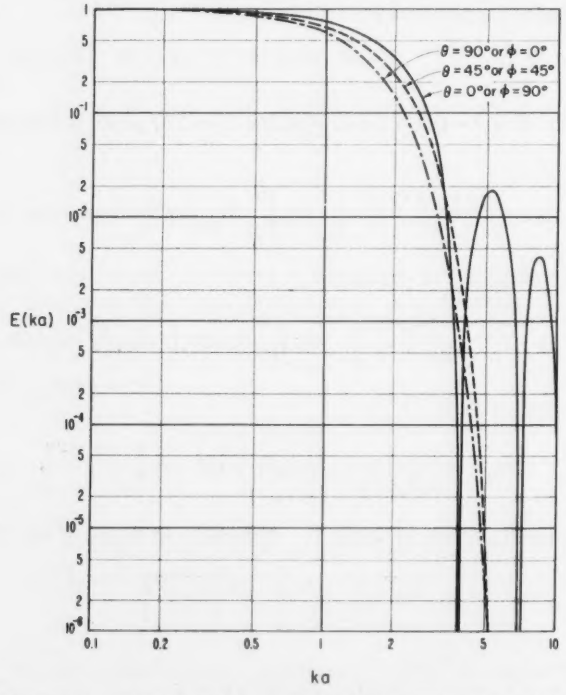


FIGURE 4. *Filter factor for a cylindrical refractometer cavity with its axis in the horizontal direction.*

θ is the angle a vector $\vec{k}$, in the xy plane makes with the vertical axis, and the angle ϕ is the angle a vector $\vec{k}$ makes with the axis of the cylinder.

It may be noted here that zeros of the functions shown in figure 2 and figure 4 can actually occur, the only requirement being that the weighting function for the refractometer decrease slowly over a small volume surrounding the center of the cavity.

The frequency spectrum of refractivity is given by [6]

$$W_n(\nu)=\int_0^\infty d\tau \cos \nu\tau <\Delta n(\vec{r}_1, t)\Delta n(\vec{r}_1, t+\tau)>. \tag{22}$$

Many measurements of this type, taken with a stationary refractometer, are available [2] Assuming (somewhat arbitrarily) that the structure of the turbulence is frozen in blobs which are carried past the refractometer by the wind with the constant velocity $\vec{\mu}$, one may write

$$<\Delta n(\vec{r}_1,t)\Delta n(\vec{r}_1, t+\tau)>=<\Delta n(\vec{r}_1,t)\Delta n(\vec{r}_1+\vec{\mu}\tau,t)>, \tag{23}$$

and then, using (1), it follows that

$$<\overline{\Delta n}(\vec{r}_1,t)\overline{\Delta n}(\vec{r}_1, t+\tau)>=<\overline{\Delta n}(\vec{r}_1,t)\overline{\Delta n}(\vec{r}_1+\vec{\mu}\tau,t)>. \tag{24}$$

It should be noted that in the above expressions ((23) and (24)) no allowance is made for the "self motion" [4] of the air, and this may be of considerable importance. Applying (24) and (5) in (22), one obtains

$$W_{na}(\nu)=\int_0^\infty d\tau \cos \nu\tau \frac{1}{8\pi^3}\int_v d^3k \exp\,[-\vec{ik}\cdot\vec{\mu\tau}]\,S(\vec{k})\,E(ka), \tag{25}$$

where $W_{na}(\nu)$ is the spectrum which is actually measured and $\nu=2\pi f$ where f is a frequency. For the case of isotropic turbulence (25) yields

$$W_{na}(\nu)=\frac{1}{4\pi\mu}\int_{\nu/\mu}^\infty dk k S(k) E(ka), \tag{26}$$

using either 1° or 2° for the filter factor.

Substituting $\nu=k\mu$ in (26) and differentiating both sides with respect to k, one obtains

$$\frac{kS(k)E(ka)}{\mu 4\pi}=-\frac{d}{dk}W_{na}(k\mu) \tag{27}$$

Norton [7] and Wheelon [4] each obtained the result

$$\frac{kS(k)}{\mu 4\pi}=-\frac{d}{dk}W_n(k\mu) \tag{28}$$

which, when combined with (27) yields

$$\frac{dW_n(k\mu)}{d(k\mu)}=\frac{1}{E(ka)}\frac{dW_{na}(k\mu)}{d(k\mu)} \tag{29}$$

which relates the slope of the actual spectrum to the slope of the measured spectrum.

3. Discussion of Results

Because no data are available for calculating the wave number spectrum, the interpretation of results will be confined to the frequency spectrum. It is well to bear in mind, however, the assumptions made (eqs (23) and (24)) to obtain these results.

Referring to figure 2 and figure 4 it is seen that $E(ka)\simeq 1$ for $ka<0.5$, so that in this range the filter factor has very little effect. If one assumes that $a=2.5$ centimeters, which corresponds

to a microwave refractometer cavity [2], it follows that $k<0.2$ per centimeter. For the constant wind velocity $\vec{\mu}$ of 5 meters per second (approximately 11 mph) this yields $f\leq 100/2\pi$ cycles per second as the range in which the frequency spectrum is unaffected by the filter factor. The available data [8] gives a spectrum accurate to frequencies of at most 10 cycles per second, and the filter factor does not alter the spectrum in this range.

If the filter factor does go to zero, as in figure 2 and figure 4, the first zero determines a definite upper limit for k (or f) for which the refractometer determined spectrum is valid. However, refractometer measurements are generally inaccurate because of noise, etc., before this limit is realized.

The need for more data to determine the weighting function and the two spectra is apparent. Even if such data were available, it is felt that one of the three forms derived here would give a satisfactory approximation.

4. References

[1] A. D. Wheelon, Radiowave scattering by tropospheric irregularities, J. Research, NBS, **63D,** 205 (1959).

[2] G. Birnbaum and H. E. Bussey, Amplitude, scale and spectrum of refractive index inhomogeneities in the first 125 meters of the atmosphere, Proc. I.R.E., **43,** 1412 (1955).

[3] M. C. Thompson, private communication.

[4] A. D. Wheelon, Relation of radio measurements to the spectrum of tropospheric dielectric fluctuations, J. Appl. Phys. **28,** 684 (1957).

[5] Tables of integral transforms, Bateman Manuscript Project, Vol. 1, (McGraw-Hill Book Co., New York, N.Y., 1954).

[6] Radio studies of atmospheric turbulence (to be published).

[7] K. A. Norton, Point-to-point relaying via the scatter mode of tropospheric propagation, I.R.E. Trans. on Commun. Systems, **CS-4,** No. 1, 39 (1956).

[8] K. A. Norton, Recent experimental evidence favoring the $\rho K_1(\rho)$ correlation function for describing the turbulence of refractivity in the troposphere and stratosphere (to be published in the Proc. of the Joint Comm. on Radio Meteorology held under the auspices of the Intern. Sci. Radio Union at New York Univ., August, 1957).

Boulder, Colo.

(Paper 64D1–40)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Conference on Arctic Communication

R. C. Kirby and C. G. Little

(Received July 23, 1959)

1. Introduction

A Conference on Arctic Communication was sponsored by the Central Radio Propagation Laboratory and held at the Boulder Laboratories, National Bureau of Standards, March 3 to 6, 1959. Approximately 275 persons attended, representing universities and research, consulting, and engineering organizations, as well as the Department of Defense and other government agencies. Foreign representation included Canada, Great Britain, SHAPE Air Defense Technical Center, and (by communication) the Norwegian Defence Research Establishment.

The objectives of the conference were to review the results of recent arctic radio research and to discuss current research and operational problems. The opening session was devoted to a review by various laboratories of their arctic research programs. Forty-six papers were then presented at four open and two classified sessions. The conference closed with an informal discussion period, during which three panels, devoted respectively to communication at vlf to lf, hf to vhf, and uhf, discussed fields requiring further research. The full proceedings of the conference will not be published, as it is expected that appropriate papers will be published in technical journals by individual authors.

2. Review Session

Dr. C. T. Elvey, Director of the Geophysical Institute, University of Alaska, reviewed the radio propagation research undertaken by the Institute. He first described studies of the scatter and reflection of radio waves at oblique and vertical incidences from the ionosphere. Field strength studies at various frequencies, including some transpolar work between Alaska and Norway, were also described. Dr. Elvey then briefly outlined the application of radio astronomical techniques to the study of ionospheric absorption and scintillation. An experiment involving the use of three Doppler-recording stations on artificial earth satellites, and a new program involving the study of radio wave interaction at frequencies of several megacycles, were also described. This latter experiment is designed to give information on the electron density and electron-collision frequency over the height range 50 to 80 km.

H. S. Marsh described the Air Force Cambridge Research Center's (AFCRC) propagation research program. A number of studies are supported at other research organizations and universities. Four stations are in operation for synoptic studies of transmissions from the oblique sounders at Fairbanks and Thule. In Alaska, ionospheric backscatter techniques are in use to determine electron-density profiles in the ionosphere. AFCRC is also sponsoring a study of low frequency propagation during polar blackouts. Another program is devoted to a study of the literature on arctic propagation, with emphasis on anomalous effects. AFCRC has an "in-house" program to measure the amplitude distribution of high frequency signals received from the Thule backscatter sounder.

A paper summarizing vhf and uhf radar studies of the aurora, in the absence of the authors, R. L. Leadabrand and A. M. Peterson, was read by R. Presnell. The paper reviewed the results of work by several organizations. Topics discussed included position of the aurora in space, strength of the echoes, fading rates, spectra, percent occurrence, polarization, wavelength dependence, and latitude dependence.

The fourth paper, by R. C. Kirby, T. N. Gautier, and A. D. Watt, was a survey of arctic propagation studies at the National Bureau of Standards. The work included NBS-sponsored and other agency programs carried out by the Central Radio Propagation Laboratory. Regular synoptic observations of vertical incidence ionospheric characteristics are made, as well as studies of high frequency field strength, atmospheric noise, very low frequency and low frequency propagation, and ionospheric and tropospheric scattering.

Operational studies of arctic communication problems by the U.S. Army Signal Research and Development Laboratory were described in a paper by C. E. Sharp and A. H. Waite. Investigation of factors affecting transmission ranges of military communication equipment in arctic regions were described.

A report on propagation phenomena observed in operating the DEW line over a 2-yr period was given by W. Donaldson of the International Telephone and Telegraph Laboratories. Performance of tropospheric scatter, ionospheric scatter, and high frequency communication systems was discussed, including the effects of unusual propagation phenomena, such as ducting at uhf, long range vhf interference, multipath, and backscatter. Propagation phenomena effecting the long range detection by L-band radar were also discussed.

A review of performance of polar region high frequency communication circuits was given by W. R. Vincent, R. L. Leadabrand, and A. M. Peterson,

524214—59——6

of Stanford Research Institute. Logs of selected circuits were analyzed for relation of propagation outage to magnetic disturbances.

The interest of the Voice of America in arctic radio propagation was described by G. Jacobs of the U.S. Information Agency. The operational requirements of the Voice in connection with its worldwide broadcasts were described, as well as methods for bypassing the northern auroral regions. Research sponsored by the USIA includes a visual aurora program and high latitude field strength studies.

Arctic communication experience in the North Atlantic region of the Airways and Air Communication Service was described by Col. W. E. Geyser, USAF. He described the radio communication facilities in the northern regions, including high frequency, low frequency, ionospheric scatter, tropospheric scatter, microwave intersite links, and long-line telephone. The operational reliability, flexibility, and limitations of each of these systems was discussed in relation to Air Force requirements. Col. Geyser outlined ultimate communication system requirements and some ideas for a minimum essential communication capability. He stressed the need for a technique which, though possibly severely restricted in information capacity and speed, would offer the highest possible order of propagation reliability and jamming resistance over ranges in excess of 3,000 miles.

Frequency-planning problems for arctic communication were discussed by G. W. Haydon of the Department of the Army. He stressed the lack of basic propagation data for frequency engineering, noting especially the requirements for statistical mapping of absorption in arctic regions and further measurements of noise from aurora. These factors could conceivably become limiting factors as receiver noise figures at vhf and uhf are significantly reduced. He also noted that prediction of high frequency signal strength and noise level were in an unsatisfactory state for arctic regions, and that the zonal concept continued to limit the usefulness of MUF predictions at high latitudes.

F. Lied, director of the Norwegian Defence Research Establishment, communicated a review of ionospheric studies in and near the auroral zone undertaken by that organization. Programs reviewed included studies of forward scatter, drift and fading, auroral reflections, absorption, *D*-region during blackouts, vlf and lf recordings, hf recordings, and sporadic-*E* ionization.

The following is the list of papers presented at the review session:

Survey of radio propagation research at the Geophysical Institute, University of Alaska, C. T. Elvey, Geophysical Institute, University of Alaska, College, Alaska.

Survey of Air Force Cambridge Research Center's current long-range propagation research program, H. S. Marsh, Air Force Cambridge Research Center, Bedford, Mass.

A summary of vhf and uhf radar studies of the aurora, R. L. Leadabrand and A. M. Peterson, Stanford Research Institute, Menlo Park, Calif.

Survey of NBS arctic propagation studies, R. C. Kirby, T. N. Gautier, and A. D. Watt, Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

Signal Corps operational studies of arctic communications problems, C. E. Sharp and A. H. Waite, U.S. Army Signal Research and Development Laboratory, Fort Monmouth, N.J.

ITT report on DEW-line observations, W. Donaldson, ITT Laboratories, Nutley, N.J.

A review of the performance of polar region communications circuits, W. R. Vincent, R. L. Leadabrand, and A. M. Peterson, Stanford Research Institute, Menlo Park, Calif.

The Voice of America's interest in arctic radio propagation, G. Jacobs, U.S. Information Agency, Washington, D.C.

Arctic experience of NORLANT AACS and systems limitations, Col. W. E. Geyser, Headquarters North Atlantic AACS Region, USAF, Westover AFB, Mass.

Frequency engineering problems in arctic communications, G. W. Haydon, Office of Chief Signal Officer, Department of the Army, Washington, D.C.

Ionospheric studies in and near the auroral zone undertaken by the Norwegian Defence Research Establishment (communication from F. Lied, director, Norwegian Defence Research Establishment, Oslo, Norway).

3. Classified Sessions

Two classified sessions were held. Representatives of Army, Navy, and Air Force described operational plans and requirements, and classified results of research programs were presented.

Requirements of the Strategic Air Command and arctic communication experience of that organization were discussed by Col. C. W. Bagstad, Headquarters, SAC, Offutt AFB, Nebr. A discussion of USAF future communication requirements was given by Col. T. N. Dickens, and a description of plans for fulfilling these requirements was given by Col. H. R. Johnson. Both represented Headquarters, USAF.

A further Air Force study of air/ground communication requirements in arctic regions for air defense systems was reported by Maj. W. Athas of the Air Defense Systems Integration Division and J. F. Roche of Lincoln Laboratory, Massachusetts Institute of Technology.

The Navy's communication requirements in polar regions were discussed by J. L. Harris, Office of the Chief of Naval Operations, Navy Department.

Army operational requirements for communication in the Arctic were described by A. Brown of the U.S. Army Signal Research and Development Laboratory at Fort Monmouth.

A study by SHAPE of ship/shore communication problems in the North Atlantic was reported by Dr. N. Knudtzon, Chief of the Communications Group, SHAPE Air Defense Technical Center, The Hague, Netherlands.

A special study of arctic vlf propagation, involving transmission over ice, was reported by A. G. McNish of the National Bureau of Standards.

The status of an experimental study of the usefulness of scatter propagation for air/ground communication was reported by T. F. Rogers on behalf of the Air Force Cambridge Research Center.

Arctic tropospheric scatter results were discussed in two presentations. A description of DEW DROP, "a link between Thule and Omaha," and its operational performance was given by Glen Mellen of the Massachusetts Institute of Technology. Tropospheric scatter in the arctic over high altitude paths was discussed by A. E. Teachman of Page Communication Engineers, Inc.

4. Open Technical Sessions

Twenty-two papers were presented during three open technical sessions. Abstracts are reproduced below, in alphabetical order by first author.

HF Field Strength Measurements in the Arctic

V. L. Agy, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

High frequency cw transmissions on 7.815 and 10.905 Mc from Bismarck, N.Dak., and Maui, Hawaii, were recorded at Anchorage and College, Alaska, using narrow-band receivers, for the period from about April 1, 1958 to December 31, 1958. In spite of the narrow bandwidth of the receivers, interference from stations on or near 10.905 Mc rules out the use of recordings at this frequency for detailed comparisons. The 7.815-Mc recordings, for the most part, however, have been satisfactory.

For the period covered by the recordings the following conclusions may be drawn:

1. The paths from Bismarck (Bismarck-Anchorage, Bismarck-College) are not notably different from a propagation standpoint. Certainly the large differences in field strength expected on the basis of an assumed sharply defined auroral absorption zone (more or less coincident with the visual auroral zone) do not occur.

2. Correlations exist between field strength variations observed over any pair of the four paths (Bismarck-Anchorage, Bismarck-College, Maui-Anchorage, and Maui-College) which are in general much larger than those observed for propagation paths entirely free of auroral zone effects.

3. Distributions of field strength values for the Bismarck paths for nighttime hours have an interdecile range of around 30 db, while for the Maui paths the interdecile range is about 15 db. This difference is attributed to auroral zone attenuation effects.

4. In addition, after correction for differences in transmitter power and path length, a median difference in field strength (Maui minus Bismarck) at either receiver of about 15 db can be ascribed to auroral zone absorption.

Abnormal Ionization in the Lower Ionosphere Associated With Cosmic-Ray Flux Enhancements

D. K. Bailey, *Page Communications Engineers, Inc., Washington, D.C.*

Abnormal ionization in the lower ionosphere associated with cosmic-ray flux enhancements is discussed mostly in terms of the great solar event of February 23, 1956. Two kinds of abnormality were recognized: *Early effects* observable in the dark hemisphere at the time of the sudden cosmic-ray enhancement, and *late effects* which began gradually and reached a maximum a few hours after the cosmic-ray enhancement. The late effects died away over a period of several days in geomagnetic latitudes above 70°, but more rapidly in lower latitudes. Both effects were more intense in higher geomagnetic latitudes. Early effects, though much the less intense of the two, appear to have been observable in lower geomagnetic latitudes than were the late effects.

The early effects are explained as a consequence of a plausible difference in composition between streams of solar particles of cosmic-ray energies and ordinary cosmic rays. The late effects are explained in terms of ionization produced in the lower ionosphere (the range of height from 30 to 110 km is studied) by the passage or stopping of solar particles—mostly protons. Such particles are shown, like the more energetic solar cosmic rays, to be arriving essentially isotropically after the maximum effects occur. A plausible extension of the observed magnetic-rigidity spectrum of the solar cosmic rays downward in rigidity by about a power of 10 would provide the protons necessary to account for the observations. The absence of auroral and significant magnetic activity accompanying the abnormal ionization is consistent with the explanations offered. Moreover, certain more recent high-latitude absorption events, unaccompanied by groundlevel cosmic-ray flux enhancements, are explicable in similar terms.

An interesting by-product of the investigation is a provisional evaluation not only of the coefficient of collisional detachment of electrons from negative ions, probably mostly O_2^-, but also of the negative-ion-to-electron ratio and the effective recombination coefficient *at night* between 30 and 110 km.

Predicting the Performance of Long Distance Tropospheric Communication Circuits

A. P. Barsis, K. A. Norton, and P. L. Rice, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

Performance of long-distance tropospheric communication circuits is predicted in terms of the prob-

ability of obtaining a specified grade of service or better for various percentages of the time. The grade of service is determined by the minimum acceptable ratio of hourly rms carrier to rms noise for the type of intelligence to be transmitted. It is shown that the prediction errors, expressed in decibels, have a standard deviation which depends upon the percentage of hours the specified grade of service is required and on the angular distance characterizing the propagation path. Finally, the possibility of reducing this standard deviation by making path-loss measurements is discussed. Thereby the reliability of the prediction might be increased.

As the prediction uncertainty was determined from analysis of a great amount of propagation data representing a variety of climatic regions, and the initial determination of basic transmission loss for any propagation path involves the use of meteorological parameters for the area concerned, the prediction method is equally applicable to arctic communication problems. Calculation schemes and graphical aids have been prepared which enable the practicing engineer to use the prediction method for any particular communication link.

Climatological Effects on Radio Propagation

B. R. Bean, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

The National Bureau of Standards is conducting extensive studies of climatic effects upon radio waves. Among these studies is the development of an exponential reference atmosphere which facilitates the estimation of refraction effects in any region for which surface weather observations are available. Worldwide maps of the mean value of the refractive index have been prepared for this purpose. In addition, a specific study has been made of radiosonde observations to determine the occurrence of ground-based radio ducts. Maximum observed incidence of ducts was determined as 13 percent in the tropics, 10 percent in the Arctic, and 5 percent in the temperate zone. Annual maximums are observed in the winter for the arctic and in the summer for the tropics.

Frequency Correlation in HF Auroral Propagation

J. W. Craig, *Lincoln Laboratory, Massachusetts Institute of Technology.*

Transmissions at spaced frequencies were made from Thule to Ipswich, Mass., to study the effect of disturbed conditions on the correlations of such signals. The data are now substantially reduced and some indication of the "correlation distance" in frequency obtained. Implications for possible new communication systems to operate under disturbed conditions will be discussed.

Performance of the Canadian Janet B Equipment in the Auroral Zone

J. H. Crysdale, *Defence Research Telecommunications Establishment, Ottawa, Ontario*

The Canadian Janet B equipment has been operated on a circuit in the auroral zone during the past year. The terminals of this circuit, which is approximately north-south and 1,000 km in length, are located at Edmonton and Yellowknife. The circuit is operated at two closely spaced frequencies in the vicinity of 40 Mc and at two closely spaced frequencies in the vicinity of 50 Mc.

One of the most serious problems encountered is auroral zone blackout associated with solar flares. During July, the circuit was not usable for at least two days because of this phenomenon. From a review of the performance of other vhf circuits, it has been found that extended blackouts are by no means infrequent.

Experience to date suggests that the meteoric signal does not predominate during substantial proportions of the time. A deeply fading signal of essentially constant or slowly varying peak amplitude is often obtained and is associated with moderate error rates. Very rapidly fluctuating signals, often of considerable amplitude and probably associated with aurora, are frequently observed. With signals of this type, the error rate can be excessive, the transmissions in many cases being garbled completely.

The analysis of circuit performance is being expedited by the use of an IBM 650 computer which is programed to obtain fairly complete burst length and error rate information from the received test message tapes.

A Proposal for Sweep-Frequency Soundings of the Ionosphere With Oblique Incidence at High Latitudes

K. Davies, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

High frequency communications (3 to 30 Mc) are still the most convenient for many purposes in the Arctic regions. In order to predict such characteristics as maximum usable frequency and received field strength, a knowledge of the mode structure of the received signal is desirable. The most convenient method of determining this mode structure is the sweep frequency pulse technique. This technique is generally used in vertical incidence ionospheric soundings; and hence, a knowledge of the relationship between vertical incidence and oblique incidence data is essential as a first step in the solution of this problem.

A most desirable feature in any oblique-incidence experiment is a vertical incidence ionosonde operating at or near the midpoint of the path. This requirement is difficult to realize in the Arctic and, as far as is known, has never been achieved. However, at present there are three vertical incidence stations

located on the west coast of Greenland which are nearly equidistant. These stations are at Thule, Godhavn, and Narsarssuak. Thule and Godhavn are about 1,900 km apart and Godhavn is within 50 km of the midpoint.

The existence of these stations thus provides a unique opportunity to investigate radio propagation in a normally inaccessible area.

VLF Aircraft Measurements

R. H. Doherty, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

Three vlf stations, NSS, NPG, and GBZ (Annapolis, Seattle, and Rugby) were monitored on a flight into the arctic region of Greenland. The flights were made in the winter, and the propagation paths were mainly nighttime paths.

Signals which had traversed the icecap were not excessively attenuated, but flights over the cap indicated a large directional effect. The directional effect has been attributed to the use of a trailing wire antenna and the large wave tilt associated with the low conductivity of the icecap. Preliminary calculations have been made in an attempt to evaluate the conductivity of the icecap. These results are compared with Dorsey's [1] values for the physical constants of ice. The values deduced from the experiment were considerably higher, indicating influence of the more highly conductive ground beneath the ice.

Utilization of Sporadic-*E* Propagation in the Arctic

R. M. Gallet, H. I. Leighton, and E. K. Smith, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

A normal procedure in hf communication in the Arctic is to decrease transmission frequency during disturbed periods inasmuch as the F_2 muf is normally depressed during such times. Actually sporadic E is more prevalent during disturbed than quiet periods and can be used to back up medium distance hf circuits (1,000 to 2,000 km) even during periods of polar blackout. However, to make the best use of sporadic E the transmission frequency needs to be increased rather than reduced.

Es data obtained from the worldwide network of ionosphere sounders indicate that sporadic E in the auroral zone tends to be most prevalent near the 100-percent isochasm (maximum occurrence of aurorae) and in the neighborhood of geomagnetic midnight. Sporadic E recorded at 24 and 48 Mc on an experimental scatter circuit between Anchorage and Barrow, Alaska, and on the College, Alaska, ionosphere sounder is used to illustrate the temporal variations of sporadic E and signal levels to be expected when the *Es* mode of transmission is used.

[1] N. E. Dorsey, Properties of ordinary water substance (Reinhold Publishing Co., New York, N.Y., 1940).

Summary of Path Loss Measurements on Tropospheric Beyond-Horizon Radio Paths

R. M. Hawekotte, *Bell Telephone Laboratories, Whippany Laboratory, Whippany, N.J.*

One of the most important factors that must be considered in the engineering of tropospheric beyond-horizon radio relay systems is the loss on the path between the transmitting and receiving antennas. Path loss measurements have been made on a number of paths ranging in length from about 60 to more than 800 mi. Frequencies ranging from about 100 Mc to more than 4,000 Mc have been used on these various paths. Some of these published results of loss on paths over land, frozen sea water, and open sea water have been reduced to 800 Mc for engineering of military communication systems in the 750- to 950-Mc band. The 800-Mc scatter loss on a path over land is in close agreement with that on a path of equal effective length over frozen sea water, but is greater than the loss on a similar path over open sea water.

200-Mc Forward Scatter Results Which May Have Implication for Arctic Communication

J. L. Heritage, *Smyth Research Associates, San Diego, Calif.*

In 1958 a survey was made of ionospheric signals received from a powerful 200-Mc transmitter. On oblique transmission paths, at geomagnetic latitude about 40°, signals were received which resembled auroral propagation in several ways.

Great-circle ionospheric forward scatter was below detectability while rapidly fading signals with 30-db signal-to-noise ratio were received at sites to which specular transmission was possible via E-layer ion columns alined with the earth's magnetic field.

These signals were named the H_E scatter because of their connection with the earth's magnetic field at E-layer height. The system sensitivity was high—the minimum transmission loss of the H_E scatter relative to free space loss was about 105 db. This figure is roughly comparable with the 40-db transmission loss reported by Dyce for amateur 220-Mc auroral transmissions at higher latitude.

The H_E scatter signals have two forms, a short burst-like form and a longer lasting form—lasting up to 20 min. Both forms exhibit a rapid fading from the onset. The repetition rate of pulsed transmission was low—the autocorrelation function was zero for adjacent pulses.

The diurnal variation of duty cycle is similar to that found for auroral transmission. During night hours H_E scatter duty cycle swamps the meteor burst duty cycle. Seasonal data is not yet available.

The above results suggest that at higher geomagnetic latitudes, systems of intermediate sensitivity might find H_E scatter a useful propagation mode. It is advocated that a number of experiments be undertaken to learn more about the frequency and latitude dependence of this form of scatter.

Results on Critical Frequency Studies for the Arctic

G. E. Hill, *AVCO Research and Advanced Development Division, Wilmington, Mass.*

Monthly medians of f_oE and f_oF_2 for January, March, and June 1957, were collected from all available Northern Hemisphere ionospheric sounding stations. The data were arranged according to G.m.t., in 3-hr intervals, plotted on polar stereographic maps, and analyzed for isopleths of the critical frequencies (eight maps per month). Instantaneous synoptic data from the IGY network, including f_oE_s, were handled similarly for 11 through 15 July 1957 (eight maps per day).

The monthly maps show the following: As expected the daytime f_oE essentially follows a cosine dependence on the solar altitude. The f_oF_2 generally decreases toward high latitudes forming a "polar low." This polar low is more extensive in winter than summer. This polar low moves around the globe at high latitudes following the sun, the winter radius being about twice the summer radius. The high daytime winter values of f_oF_2 in auroral latitudes are related to the position of the polar low. Evidence of solar tidal effects is found (the polar low becomes a complex of three lows).

The synoptic series in July show that the features evident on a monthly average are present on a daily basis also. The strong regularity of the E-layer appears daily as well as monthly. The daily features of f_oF_2, however, show some important departures from the monthly median variations. These departures, or disturbances, have characteristic velocities ranging from about zero to several hundred m/sec. There exist preferred areas of development, but these areas are not permanent. Depressions occur at all times of the day, frequently about noon. The occurrence of sporadic E in temperate latitudes is mainly in a broad belt nearly encircling the earth.

A minimum is found at early morning longitudes. At night E_s is found in northern latitudes. During the five day period wide departures from this simplified picture for E_s occur. There appears to be some correspondence between patterns of f_oF_2 and E_s occurrence.

The effect of the ionospheric general circulation is discussed, although no specific circulation model is presented.

Planning Intersite Paths for Communications Facilities

R. S. Kirby, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

In selecting terminal locations for long distance communication facilities, such as scatter and hf circuits, one of the primary site conditions is to have good intersite paths to the communication relay centers. In remote areas of the world this intersiting is normally accomplished with a microwave relay operating in the 2,000- to 8,000-Mc range. Often the sites are situated in such a way that line of site paths do not exist between the two ends. With low-powered relay equipment currently in use, this requires the siting of one or more relays at remote intermediate locations.

With higher powered relay transmitters, up to 2 kw at 8,000 Mc, the siting requirements for intersite paths can be relaxed and nonoptical paths can conceivably be used. This paper presents a study of the propagation problems associated with transmitting high-power broad-band shf signals over the horizon using diffraction and scatter modes of propagation. Particular attention is given to special meteorological conditions prevalent in the Arctic.

HF Radio Wave Absorption in Northern Latitudes

H. Leinbach and G. C. Reid, *Geophysical Institute, University of Alaska, College, Alaska.*

Anomalous absorption in the lower ionosphere in the arctic regions can be classified into two types: (*a*) aurorally-associated absorption (type II) and (*b*) absorption arising from ionization of the lower ionosphere by charged particles from solar flares (type III).

Aurorally-associated absorption (type II) is strongest at the southern edge of the auroral zone, i.e., in the vicinity of 65° geomagnetic latitude. Type II absorption reaches a maximum at this latitude during the post-breakup aurora (early morning hours, local time), often reaching hourly average values in excess of 2 or 3 db at 27 Mc. A 5-Mc wave passing once through the absorbing region at moderate obliquity could thus suffer upwards of 200-db attenuation, so that hf communication would become impossible using normally attainable powers.

Type III absorption is much less frequent, but is of considerable importance since complete absorption of hf waves may occur over the entire polar region north of 65° geomagnetic latitude. Furthermore, the absorption may remain heavy for periods of up to four days. Six to eight type III events per year may be expected during the sunspot maximum years.

The above statements are based primarily on IGY absorption data obtained in Alaska and Greenland using the cosmic noise technique at 27.6 Mc.

Radio Star Scintillation Studies at College, Alaska

C. G. Little,[2] *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

The methods used in a study of the amplitude and angular scintillations of the Cygnus and Cassiopeia

[2] Formerly of Geophysical Institute, University of Alaska.

sources are briefly described. The results are presented in diagrams giving:

(1) The probability of occurrence of five different levels of scintillation activity.
(2) The probability, for each level of activity, of observing a given increase or decrease of signal strength.
(3) The probability, for each level of activity, of observing angular scintillations greater than a given value.
(4) The solar-time variation of the mean fractional deviation of power, $\overline{\frac{|\Delta P|}{P}}$, for each source.
(5) The elevation angle dependence of $\overline{\frac{|\Delta P|}{P}}$ for each source.

The observations have led to the following conclusions relating to the College data:

a. The scintillation activity during 1957–58 was some 20 times more intense than in England during 1949–51.

b. The solar-time variation of scintillation activity is very much less than at temperate latitudes.

c. The scintillation amplitudes were largely independent of the zenith angle of the source.

d. The scintillation activity of the Cassiopeia source was significantly less than for the Cygnus source, owing to its large angular size.

e. The ionospheric blobs responsible for the scintillations are elongated along the geomagnetic field lines by a factor of about 6. The 0.5 correlation distance transverse to the magnetic field is of the order 300 m.

f. The rms deviation in electron density in the irregularities is of the order 500 electrons/cm^3 (or less) during quiet periods and is approximately 5,000 electrons/cm^3 during disturbed conditions. Values of 20,000 electrons/cm^3 would be adequate to explain the very occasional (less than 1% of the time) disappearance of the source on the interferometer recordings.

Application of Data from the Vertical Incidence Ionospheric Network to Auroral Zone and Polar Cap Communication Problems

M. E. Nason, *National Bureau of Standards, North Pacific Warning Service, Anchorage, Alaska.*

A network of eight NBS associated vertical sounding stations maintains a continuous schedule of ionospheric observations at auroral zone and polar cap locations ranging from Reykjavik, Iceland, to Barrow, Alaska. These data are primarily useful for long-range predictions of maximum useful frequencies. However, by telegraphic reports they can be used in the day-to-day forecasting of short term fluctuations of the MUF over important radio paths. Auroral zone and polar cap absorption effects are observable and of potential use in assessing the degree of probable disturbance.

Backscattering and Propagation of Radio Waves

L. Owren, H. Bates, and R. Hunsacker, *Geophysica Institute, University of Alaska, College, Alaska.*

Oblique incidence sweep-frequency backscatter sounding of the arctic ionosphere over the frequency range 1 to 25 Mc has shown a prevalence of anomalous echoes. A direct F-layer echo is observed regularly and is associated with electron density irregularities at actual heights between 350 and 600 km. Both the $1F$ (direct) and $2F$ (groundscatter) echoes received at College, Alaska, from the north usually connect to the extraordinary branch of the vertical incidence trace while the $2F$ echoes from the south appear to connect to the ordinary branch. $2F$ echoes from the north and south indicate that at times the reflecting layer is tilted 2° to 3°.

The relative signal strength of fixed frequency pulse transmissions on 12, 18, and 30 Mc from College, Alaska, recorded at Kiruna, Sweden, have been compared with simultaneous groundscatter echoes received at College from the polar region. Analysis for December 1958 shows that the 12- and 18-Mc signals are propagated by two and three-hop modes as well as a one-hop tilted layer mode. The two-hop mode is mainly a daytime mode, Alaska time, while the dominant three-hop mode occurs both day and night. Minimum received signal and polar groundscatter is observed at the time of maximum D-layer absorption for the outgoing radiation.

The research program is supported by the Electronics Research Directorate, Air Force Cambridge Research Center.

Theory of Spread *F*

Jacques Reneau, *Cornell University*

On the assumption that aspect sensitivity plays a major role in the energy backscattered from ionized irregularities elongated along the earth's magnetic field, a formula for the delay time associated with a ray from the ionospheric sounder to the position of perpendicularity in the F-region, with allowance for refraction, has been derived.

In the hope that this mechanism may explain the phenomenon of spread F, computations of delay time versus frequency were carried out for several stations located at various magnetic latitudes, and the results compared to observed ionograms. It is concluded that the model does not explain arctic spread F but clarifies the main features of the equatorial spread F.

A Proposed Technique for Improving the Intelligibility of Voice Radio Transmissions in the Presence of Auroral Flutter Fading

O. G. Villard, Jr., and K. C. Yeh, *Radio Propagation Laboratory, Stanford University, Stanford, Calif.*

It would appear to be possible to make a significant improvement in the intelligibility of voice radio transmissions garbled by auroral flutter fading, by

transmitting the voice at a higher-than-normal rate, followed by an equal slow-down at the receiving end of the circuit. This has the effect of reducing the effective speed at which the individual frequency components of the voice signal are made to fade as a consequence of transmission via a disturbed ionosphere. A slow-down in fading speed can be expected to improve intelligibility because present evidence suggests that the voice quality of amplitude-modulated transmissions falls off very rapidly when the average carrier fading speed becomes higher than roughly 10 cps.

Since the increased transmission speed requires an increase in bandwidth and a consequent reduction in received signal-to-noise ratio, the method will be useful only when adequate signal-to-noise ratios are initially present. The required equipment is relatively simple and adaptable to existing communication systems.

Some Sporadic Phenomena of the Auroral Zone

E. S. Warren, *Defence Research Telecommunications Establishment, Shirley Bay, Ontario*

Some effects of the sporadic ionization revealed by ionograms in the auroral zone are reviewed. The oblique incidence equivalents of spread F, polar spurs, and forked records are illustrated by oblique sounding records. The influence of these phenomena on hf point-to-point communication is discussed. Suggestions are made for the improvement of long-term predictions, in particular for the calculation of the MUF and the LUF and for handling the statistical variations. For the operators problem it is suggested that information from a few oblique-incidence sounders be disseminated as it is produced.

Observations on Some Low Frequency Propagation Paths in Arctic Areas

A. D. Watt, E. L. Maxwell, and E. H. Whelan, *Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.*

The very low ground conductivities encountered in arctic area, and the particular ionospheric conditions prevailing at high latitude, can lead to rather unusual radiation and propagation conditions. In order to determine the magnitude of these effects, field intensities from transmitters located in the Labrador and Greenland areas were measured both on the surface of the earth and during several aircraft flights over this area.

Under conditions where the initial portion of the propagation path is across icecap or permafrost, the attenuation observed is very great, and when the propagation path extends out over sea water, the field intensity recovery taking place after the coastline is crossed is very marked. Estimates of skywave field intensity appear to agree with the observed results provided the radiated field pattern is suitably modified by the antenna cutback factor which accounts for the presence of a finitely conducting curved earth. These vertical patterns based on work by Wait, along with the field intensity flight data, indicate that the siting of low frequency stations several miles or more inland in arctic regions may cause a great increase in total transmission path loss.

Initial Investigation of Techniques for Prediction of Short-Term f_0F_2 Variations in the Higher Latitudes

J. M. Weldon, *National Bureau of Standards, North Atlantic Warning Service, Fort Belvoir, Va.*

This paper discusses initial attempts by the North Atlantic Radio Warning Service to develop techniques for predicting high frequency radio propagation conditions over specific high latitude transmission paths. A description of the radio forecasting methods used in the general area forecast service currently provided is presented. It is suggested that while the service now offered is quite reliable over the general North Atlantic Path for which it is primarily intended, i.e., New York to London, Washington to Paris, etc., there is need for the inclusion of a specific path forecasting mechanism designed for operational use in the higher latitude circuits which are more susceptible to geomagnetic disturbance. The problem is approached by studying the ionosphere along the particular paths with the aim of finding useful relationships that can be applied toward predicting radio quality and proper frequency ranges to employ for daytime, nighttime, and transition periods. Correlative studies between F_2 layer critical frequencies at probable control points of several paths and solar activity indices, magnetic indices, and radio quality data are described generally and conclusions are drawn based on the results. Large negative deviations are readily associated with predictable magnetically active periods. However, to date few encouraging relationships have been found involving magnetically quiet period f_0F_2 variations. Future experiments are suggested to test further the feasibility of specific path forecasting.

Boulder, Colo. (Paper 64D1–41)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Tropospheric Scatter Propagation and Atmospheric Circulations

William F. Moler[1] and D. B. Holden[1]

(August 10, 1959)

Transhorizon vhf and uhf fields exhibit deep fades or large signal enhancements of several hours' duration, as the propagation mechanism alternates between partial reflection and scattering caused by turbulent dielectric fluctuations in the atmosphere. Such alternations occur when strong refractive layers develop below 3,000 ft. Surface wind streamline analyses show that mesoscale centers of convergence or divergence cause local redistribution of refractive layering, tending to produce the change from one mechanism to the other.

Current scattering theory and the empirical findings of others are examined to determine the gross meteorological factors that influence changes in scattered fields. The two variables in the turbulent scattering coefficients, the scattering angle and the intensity of dielectric fluctuations at high wave numbers, are found to be dependent upon the refractive layering and the thermal stability of the airmass. It has been shown elsewhere that refractivity and stability are principally functions of the vertical velocity in the atmosphere. It is shown here that the direction and relative magnitude of the vertical velocity can be inferred from the upper-tropospheric wind velocity divergence. Received scattered signals are found to be well correlated with computed velocity divergence.

It is suggested that the variations of scattered signal level or range can be predicted in a routine manner by regular meteorological personnel using ordinarily available meteorological data.

1. Introduction

Future Navy communication networks and early warning radar may depend to a great extent upon the tropospheric scatter mode of radio propagation. The development of high-gain low-noise receivers and statistical target identification techniques makes the vhf and uhf region of the radio spectrum useful for ranges well beyond the radio horizon.

The signal in the far transhorizon region is characterized by a rapidly fluctuating power amplitude superimposed on a more slowly changing average signal level. The rate of fast fading depends on distance and radiofrequency, indicating the existence of uncorrelated multipath transmission from numerous moving scattering centers. The slow fading has seasonal, average diurnal, and random day-to-day or hourly components. Seasonal and average diurnal fluctuations are amenable to statistical analysis and are predictable within relatively narrow limits) [1].[2] The random daily or hourly variations in signal level have been less well understood. These are the fluctuations in signal level (or range) in which the Navy may be most interested.

The best current theories explain the scatter propagation mechanism by making use of scattering coefficients which are functionally related to the gross refractivity of the lower atmosphere and the intensity of turbulent dielectric fluctuations at radiofrequency-dependent scale sizes. The scattering coefficients are difficult to measure and, at the present time, are unpredictable. It is possible to predict some aspects of the atmosphere's refractive structure which, from elementary considerations of theory, are most likely to control the magnitude of the scattering coefficients and hence the transhorizon signal strength.

This report indicates the meteorological factors which appear to control the electromagnetic scattering mechanism and suggests an operational method for predicting the direction and relative magnitude of signal level change.

2. Discussion of Theoretical and Empirical Relationships

The theories proposed to explain the large other-than-diffraction transhorizon fields are based on two propagation mechanisms: (1) partial reflection of incident radiation by horizontally stratified dielectric layers [2, 3, 4], and (2) scattering by turbulent discontinuities in the dielectric structure of the atmosphere [5, 6, 7]. The phenomenon of partial reflection has been recognized and studied for over ten years. In the microwave region, for usual refractive gradients, the effect of layers decreases exponentially with height, so that reflections of this sort become unimportant for layers above about 3,000 ft. More recently, Bauer, Friis, et al. [8, 9] theorized that energy may be partially reflected in much the same manner from thin, highly refractive sublamina within the region illuminated by both antennas. Refractometer measurements made during convective atmospheric conditions by Navy Electronics Laboratory (NEL) failed to show the existence of such layering at times when transhorizon radio transmission was normal. It is likely that, under conditions of a well-established elevated layer, both propagation mechanisms may exist concurrently [10], but that in the absence of strong relatively low elevated layers, the turbulent scattering mechanism is dominant.

[1] U.S. Navy Electronics Laboratory, San Diego 52, Calif.
[2] Figures in brackets indicate the literature references at the end of this paper.

524214—59——7

The turbulent scattering theories, following the original work of Booker and Gordon [5], postulate in common an equation of the general form:

$$Pr \propto \frac{1}{\sin^n \frac{\theta}{2}} F(K). \quad (1)$$

Pr=received power,
θ=intersection angle of rays from transmitter and receiver,

$$K=\frac{4\pi \sin \frac{\theta}{2}}{\lambda},$$

λ=radio wavelength,
$F(K)$=function of dielectric fluctuation intensity at wave number, k, and
n=exponent, dependent on form of F.

By incorporating the modern theories of turbulence put forth by Obukhoff, Kolmogoroff, and others with radio scattering theory, Batchelor [7] found that, for an isotropic radiator, n has the value 11/3. For propagation with narrow beam antennas, the exponent becomes 14/3. It should be noted that for grazing propagation and ducting gradients, expression (1) fails; propagation is line-of-sight and, disregarding focusing effects, power is maximized.

2.1. Scattering Angle

The strong dependence of signal level upon scattering angle (mean refractive gradients) is demonstrated in figure 1. The abscissa is the signal level in decibels above that received in a standard atmosphere (-12 N units/10^3 ft) for ratios of the standard atmosphere scattering angle to the actual scattering angle (ordinate on the left) given by the N gradients (along the right of the graph). It is noteworthy that with high signal levels and near-ducting gradients, deep fades usually associated with other modes of propagation can occur with very slight changes in the low level refractive gradients.

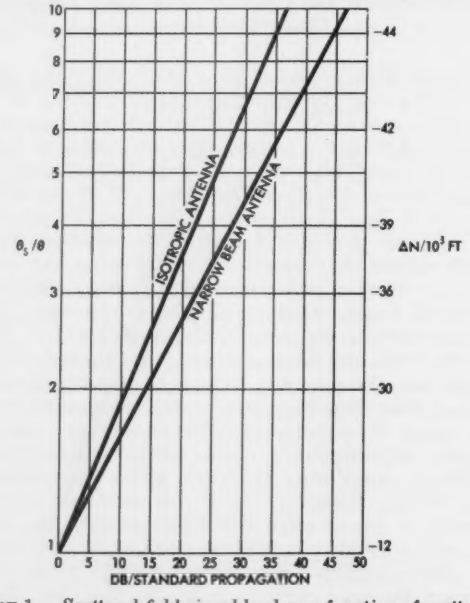


FIGURE 1. *Scattered field signal level as a function of scattering angle and refractive gradients in a homogeneous refractive atmosphere normalized to standard gradient.*

A method for determining the scattering angle from the refractive index profiles over the scatter propagation link is outlined in appendix A. The method is generalized for a multilayered atmosphere, the layers having a limited degree of horizontal homogeneity.

Other investigators concerned with developing methods for predicting transhorizon fields through use of meterorological parameters have found correlation between scattered fields and low-level refractive gradients. Bean [1], reasoning that lower tropospheric refractive gradients are directly related to the surface value of refractive index, found statistical correlation between transhorizon fields and the surface refractivity. Misme [11], using a weighted average tropospheric refractivity, found correlation between this "useful gradient" and the received signal.

Time-averaged low-level refractive gradients have been shown by Moler and Arvola [12], to be principally a function of the large-scale vertical motions in the atmosphere. Later in this report, computations of convergence at the 300-millibar (mb) level and the implied subsidence (see app. B) throughout the troposphere are compared with the development and descent of a strong refractive layer and the resultant change in scattering angle and radio signal strength.

Low-level gradients may also be temporarily altered with the development of minor eddies in the surface wind flow. Occasionally, during the NEL scatter propagation experiments, radio signal levels and characteristics would indicate a change in the propagation mechanism from scattering to partial reflections. Subsequent analysis of the small-scale wind circulations showed that local regions of convergence (divergence) which develop within a relatively homogeneous large-scale weather regime cause raising (lowering) of refractive layers and, consequently, a change from one dominant mode of propagation to another. Examples of this will be given in a later section of this report.

2.2 Turbulent Dielectric Fluctuations

Little is known about the variation of intensity of turbulent dielectric fluctuations at wave numbers important for microwave scattering as a function of the meterorological situation. Recent work by Gossard [13] indicated that the refractive index spectral intensity at high wavenumbers in the inertial subrange of turbulence is only weakly dependent upon atmospheric stability or refractive gradients. From this work, it may be concluded that variations

in signal level result almost entirely from refractive changes in scattering angle or that there are thin zones of quite variable turbulent intensity which superimpose an amplitude modulation upon a low average signal.

Batchelor [7] suggested that the fluctuations of refractive index may be due to turbulent convection in the presence of a spatial gradient of mean refractive index. If this suggestion is correct, it seems likely that the greatest intensity of dielectric fluctuation would occur along the lower boundary of elevated stable layers (refractive layers) where convective energy from the lower, less stable region is absorbed and dissipated.

Based on empirical evidence, Misme [11] found that "for equal refractive index gradients, the atmosphere which is characterized by the greatest stability permits the highest fields to be received from a great distance." By "equal gradients", he meant the refractivity which determines the scattering angle.

Gordon [14], using assumed values of fluctuation intensity, concluded that the contribution of scattered energy from a single turbulent layer may be greater than that from the remaining common volume. The theories of partial reflection [2, 3, 4, 8, 9] require the presence of highly refractive substrata within a mean refractive gradient with greatest radio fields associated with intense elevated refractive layers.

All evidence indicates that departures of radio fields from a low ambient level can best be forecast indirectly through a prediction of the atmospheric thermal stability. It is well known that, in general, atmospheric stability is largely a function of the vertical components of the winds. At the onset of subsidence, refractive layers develop; continuous subsidence causes lowering and intensification of the layers and increases the thermal stability of the airmass. Rising motion tends first to elevate the layers and later to cause their complete dissipation.

The behavior of radio signal levels under the influence of positive as opposed to negative vertical motion will be shown in a later section. The method for deducing the sign and relative magnitude of the vertical wind component from upper tropospheric constant pressure maps will be found in appendix B.

3. Operations

In January 1958, NEL activated three transhorizon radio links which were in periodic operation until mid-April. Stations at Santa Barbara (190 statute miles) and at Point Mugu (144 statute miles) transmitted 9,380 Mc (*X*-band) and 1,365 Mc (*L*-band) pulsed signals. A third station on San Clemente Island (78 statute miles) transmitted at 9,380 Mc only. All receivers were located on Point Loma, San Diego. Figure 2 is a map of the southern California operating area. Signal levels used in this report were measured over 1-min intervals on integrating equipment and recorded in decibels with respect to 1 milliwatt.

FIGURE 2. *Navy Electronics Laboratory scatter propagation experiment operations area.*

Strong low-level elevated refractive layers are a dominant feature along the southern California coast and frequently cause ducting or abnormally strong signal enhancement from internal reflections when the layers are lower than about 3,000 ft [3, 4]. In order to avoid such conditions, the links were placed in operation only when there was a reaonable assurance that refractive layers would be absent or at a high elevation. Synoptic experience has shown that airmasses are most unstable and thermal layering absent when under the influence of a strong cyclonic circulation through most of the troposphere. For this reason, the bulk of data was taken during periods when a deep low pressure system at the 500-mb level (18,000 ft) was near or over the radio link.

Radio link operations were discontinued whenever signal characteristics indicated that a propagation mode other than scattering was dominant and likely to continue. Unfortunately, as a result very little data were taken during periods of high signal level and conditions of airmass subsidence.

The southern California sectional maps were constructed from hourly airways teletype reports and from weather logs of two aircraft carriers which were operating in coastal waters. The surface synoptic and 500-mb charts are copies of National Weather Analysis Center maps transmitted over the facsimile network. Calculations of vorticity advection at the 300-mb levels were made from maps plotted at NEL, using data from the Daily Bulletin of Northern Hemisphere Data Tabulations published by the U.S. Weather Bureau.

4. Measurements

4.1. Signal Level and Refractive Layering

The constant pressure contours for the 300-mb pressure surface for 1600 P.s.t., March 16 to 19, are shown in figure 3. The field of relative vorticity (dashed lines) indicates positive vorticity advection (horizontal divergence) east of the upper level trough and negative vorticity advection (horizontal convergence) west of the trough. From continuity con-

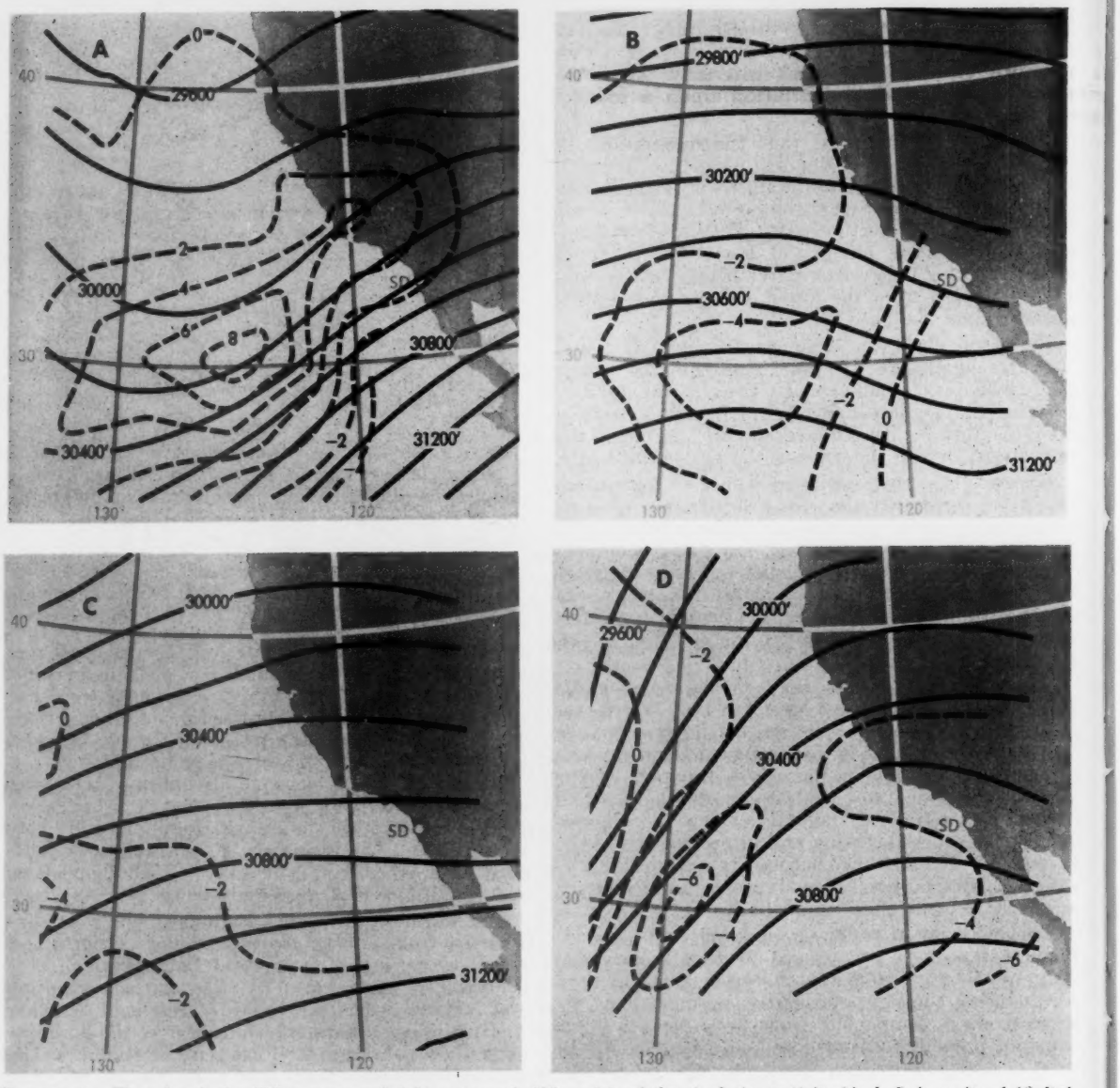


FIGURE 3. *Three hundred millibar pressure-height contours (solid) and isopleths of relative vorticity (dashed) in units of* $10^{-5}t^{-1}$ *at 1600 P.s.t. for (a) 16 March (b) 17 March (c) 18 March (d) 19 March.*

siderations (app. B), positive vorticity advection implies negative vertical velocities. Figure 4 shows $\nabla \cdot V$ 300 mb as a function of time compared with a series of refractive index profiles derived from the Los Angeles radiosonde observations. The functional relationship between the scattering angle θ and the refractive gradients has been derived in appendix A. Figure 5 shows the scattering angle for the grazing ray and height of the common volume for the Santa Barbara link during the 3-day period. The upper curve shows the signal level as a function of the scattering angle normalized for a standard atmosphere. The second curve is the radio field strength for the Santa Barbara X-band link for the period 0800 to 1600 P.s.t., 19 March. Radio data were not taken the previous days.

On March 26 there was an anomalous 20-db enhancement of signal level during a 3-hr period near noon that appeared unexplainable on the basis of radiosonde data or by vertical velocities inferred from the upper-troposphere divergence field. The refractive index profiles reduced from Los Angeles radiosonde observations for 0400 P.s.t. and 1600 P.s.t. show an elevated layer near 4,000 ft (fig. 6)

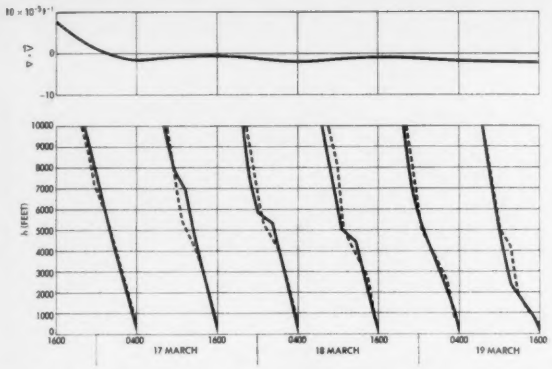


FIGURE 4. *Horizontal velocity divergence at the 300-mb pressure level (upper graph) and Los Angeles and San Diego refractive index profiles for the period 16 to 19 March 1958.*

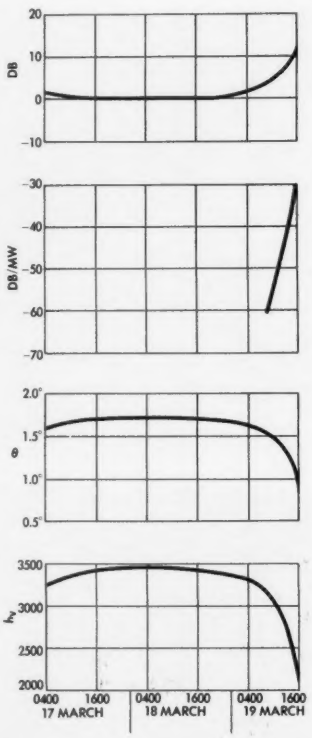


FIGURE 5. *Upper curve indicates power depending upon scattering angle, θ, for the period 17 to 19 March 1958. Second curve shows the received power in decibels per milliwatt for Santa Barbara X band. The lower two curves show the height of the scattering volume, h_v, and the scattering angle, θ, during the same period.*

and computations of $\nabla \cdot V$ made at 300 mb (upper curve fig. 6) indicate a continuous reduction in stability taking place. A series of isobaric and streamline southern California sectional charts were analyzed at hourly intervals for the day. Figure 7 shows the map series and includes the large-scale

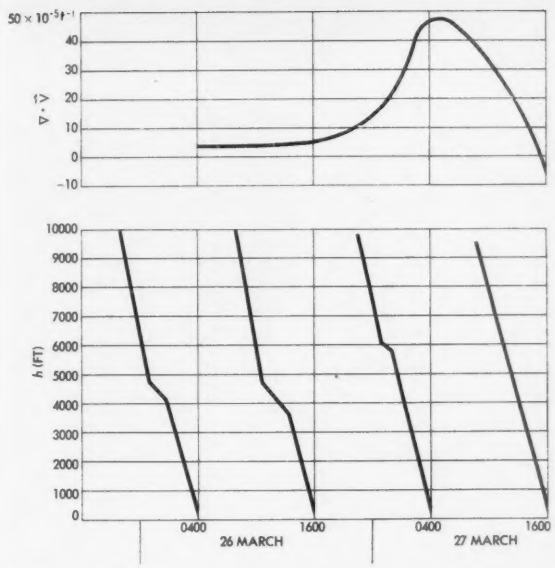


FIGURE 6. *Refractive index profiles reduced from Los Angeles rabiosonde and (upper curve) 300-mb velocity divergence for 26 to 27 March 1958.*

synoptic chart for 1000 P.s.t. Figure 8 is the Santa Barbara *X*-band signal for the day. From figure 7, at 0900 P.s.t., a cyclonic eddy with two centers northwest of San Diego lay across the Santa Barbara radio link. Streamlines indicate convergence and consequent upwelling of the surface air in the convergent areas. By 1100 P.s.t. the cyclonic eddy had weakened, due to the onset of a sea breeze and subsequent low-level divergence and subsidence of air over the Catalina channel. The 1200 P.s.t. map shows only a weak remnant of a zone of convergence south of Point Mugu and signals have risen 20 db. By 1300 P.s.t. offshore convergence had ceased entirely and signals had reached a peak level for the day shortly before 1300 P.s.t. The 1300 P.s.t. map indicates a strengthening of the pressure trough north of Point Mugu with a tendency for the trough to extend seaward. Signal levels started to fall after 1300 P.s.t. By 1500 P.s.t., the convergence zone had undergone strong redevelopment across the Santa Barbara link and signal levels had fallen about 22 db from the peak level.

4.2. Signal Level and Upper-Tropospheric Divergence

During the period 26–28 March inclusive, the radio links were in operation from 0800 P.s.t. to 1600 P.s.t. daily. The weather regime was dominated by an upper-level cyclone which developed off the northwest coast of California and moved southeastward across the radio link. A series of 500-mb height contour charts drawn at 12-hr intervals during

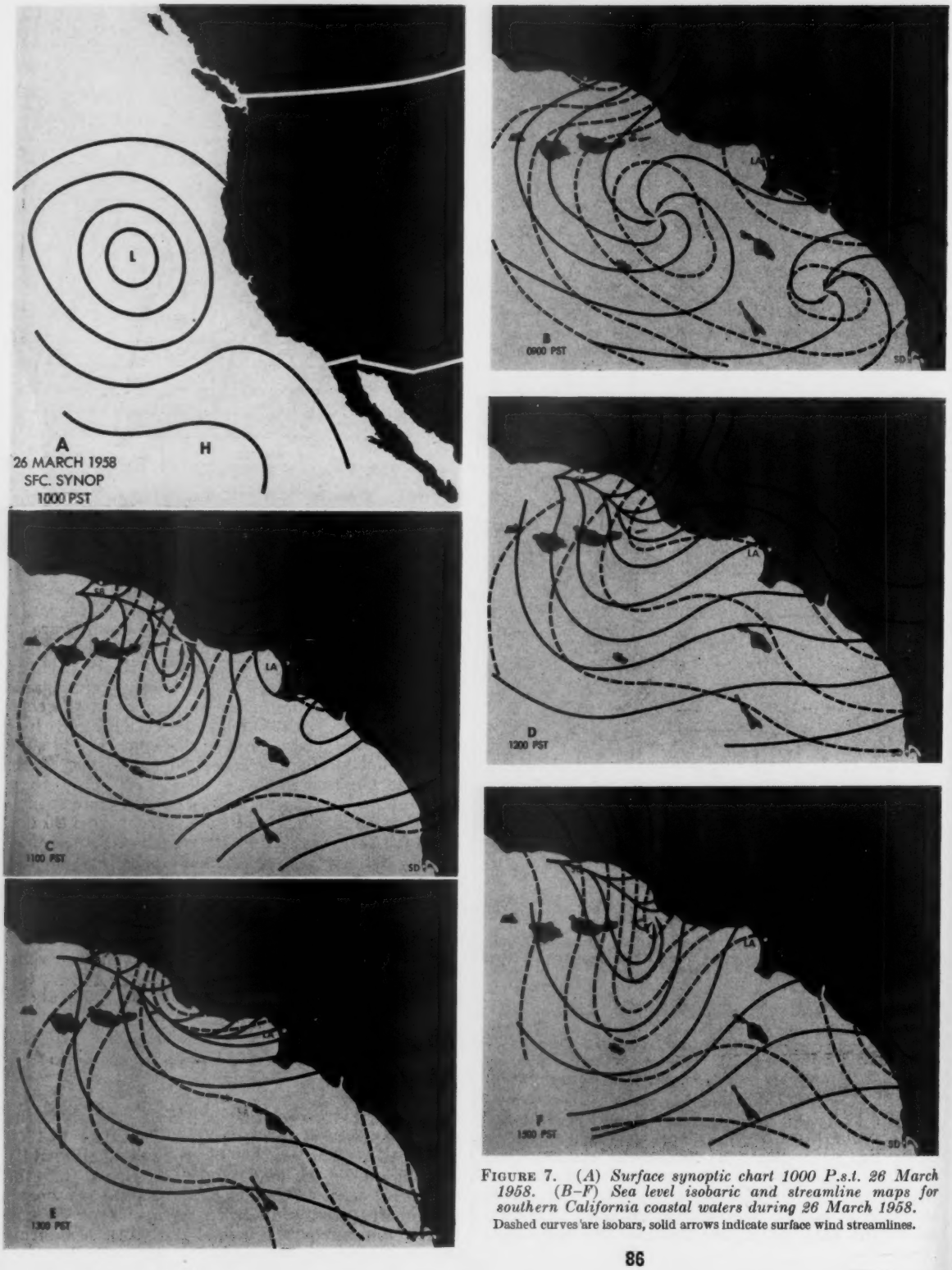


FIGURE 7. *(A) Surface synoptic chart 1000 P.s.t. 26 March 1958. (B–F) Sea level isobaric and streamline maps for southern California coastal waters during 26 March 1958.*
Dashed curves are isobars, solid arrows indicate surface wind streamlines.

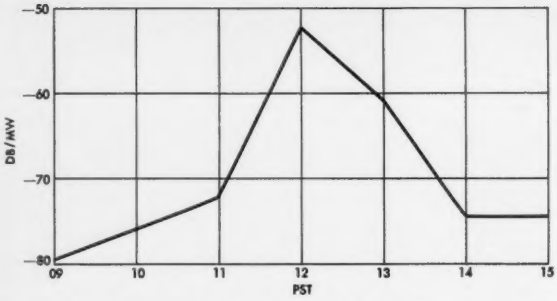


FIGURE 8. *Santa Barbara X-band signal level for 26 March 1958.*

the period are shown in figure 9. The isopleths are contours of the height of the 500-mb pressure surface drawn for 200-ft intervals. The velocity divergence along the streamlines 5° upstream from the operating area was computed for each map during the period. Figure 10 is a graph of the velocity divergence at 500 mb and average signal level received from the Santa Barbara *L*-band and Point Mugu *X*-band transmitters, both plotted as a function of time. The signal levels are average values for the periods 0800–1000 P.s.t., and 1400 to 1600 P.s.t. plotted at 0900 and 1500 P.s.t., respectively, and connected with a smooth curve. Solid portions of the curve are for hours of radio transmission. The dashed curves are extrapolations of the solid curve slopes. Unfortunately, periods of radio silence were longer than transmission periods so that one can only assume that the signal strength curves are reasonably close to those extrapolated. Although it would be impossible to predict absolute signal level, the correlation between the sign and relative magnitudes of the time derivatives of the curves is good.

Computations of vorticity advection were made at the 500-mb pressure level for all days when there was a reasonably good sample of radio data and at the 300-mb level for periods of particular meteorological interest. Signal levels for the 2-hr period nearest map time were averaged for the conditions of positive and negative vorticity advection (positive and negative vertical velocities). Table 1 gives these

TABLE 1. *Average values of signal level for conditions of positive and negative vertical wind velocities*

Pt. Magu and Santa Barbara transmitters	w>0		w<0	
	500 mb	300 mb	500 mb	300 mb
B_s	−69.5	−71.0	−62.1	−63.8
BL	−67.0	−70.0	−57.6	−60.0
M_s	−68.4	−67.1	−58.4	−61.0
ML	−66.0	−71.0	−62.3	−64.4

average values of signal under the two conditions from a sample size of 33 at the 500-mb level and 18 at the 300-mb level. Both the *X*- and *L*-band signal levels are significantly higher under conditions of subsidence with no apparent frequency dependency to be noted.

5. Results and Conclusions

Other surface streamline charts and upper-tropospheric synoptic series were analyzed, but to present the results here would be repetitious and redundant. All concurrent meteorological and radio signal level analyses showed essentially the same features as indicated in the previous sections.

During the NEL experiment, sea level values of refractive index were essentially invariant with time, yet signal levels ranged over approximately 60 db. It seems clear that for over-water paths, short-term propagation characteristics are independent of surface refractivity.

Nonducting low-level refractive layers influence transhorizon fields in two ways: (1) refraction and subsequent decrease of scattering angle, and (2) partial reflection at the interfaces of superrefractive layers as the layers approach critical reflection heights and intensities. The second effect dominates for layers below about 2,000 ft. For example, if after Batchelor [7], a dielectric fluctuation spectrum of the form $E(K) \propto K^{-\frac{5}{3}}$ is accepted, then for narrow-beam propagation, the received power dependency upon scattering angle would be $Pr \propto \theta/2^{-\frac{1}{3}4}$. Assuming standard propagation early in the morning of March 19, the Los Angeles sounding for 1600 P.s.t. March 19 (fig. 4) would indicate a lessening of scattering angle sufficient to cause a 12.7-db rise in signal level (fig. 5). The actual rise in signal was approximately 35 db. In order to account for such a signal rise by refraction, the scattering angle would have had to decrease by nearly one-sixth. It is unlikely that the overwater refractive gradients were sufficiently greater than those shown by the Los Angeles radiosonde to account for this much refraction. One is forced to conclude that the partial reflection mechanism becomes more important for transhorizon propagation than scattering when relatively low layers are present. This suggests that the refractive index climatology charts for oceanic areas [15] could be useful for determining the distribution of dominant modes of transhorizon propagation.

The short term prediction of change of radio range for oceanic areas becomes greatly simplified if the problem can be looked upon as one concerning only changes in thermal stability. Except for coastal regions where continental air may undergo rapid modification, refractive layering and thermal stability over the oceans are principally functions of vertical wind velocities. Techniques for computing exactly or making qualitative estimates of the vertical wind velocities similar to the one outlined in appendix B are familiar to meteorologists. It is likely that satisfactory estimates of radio scatter propagation ranges can be made indirectly from regularly available weather charts.

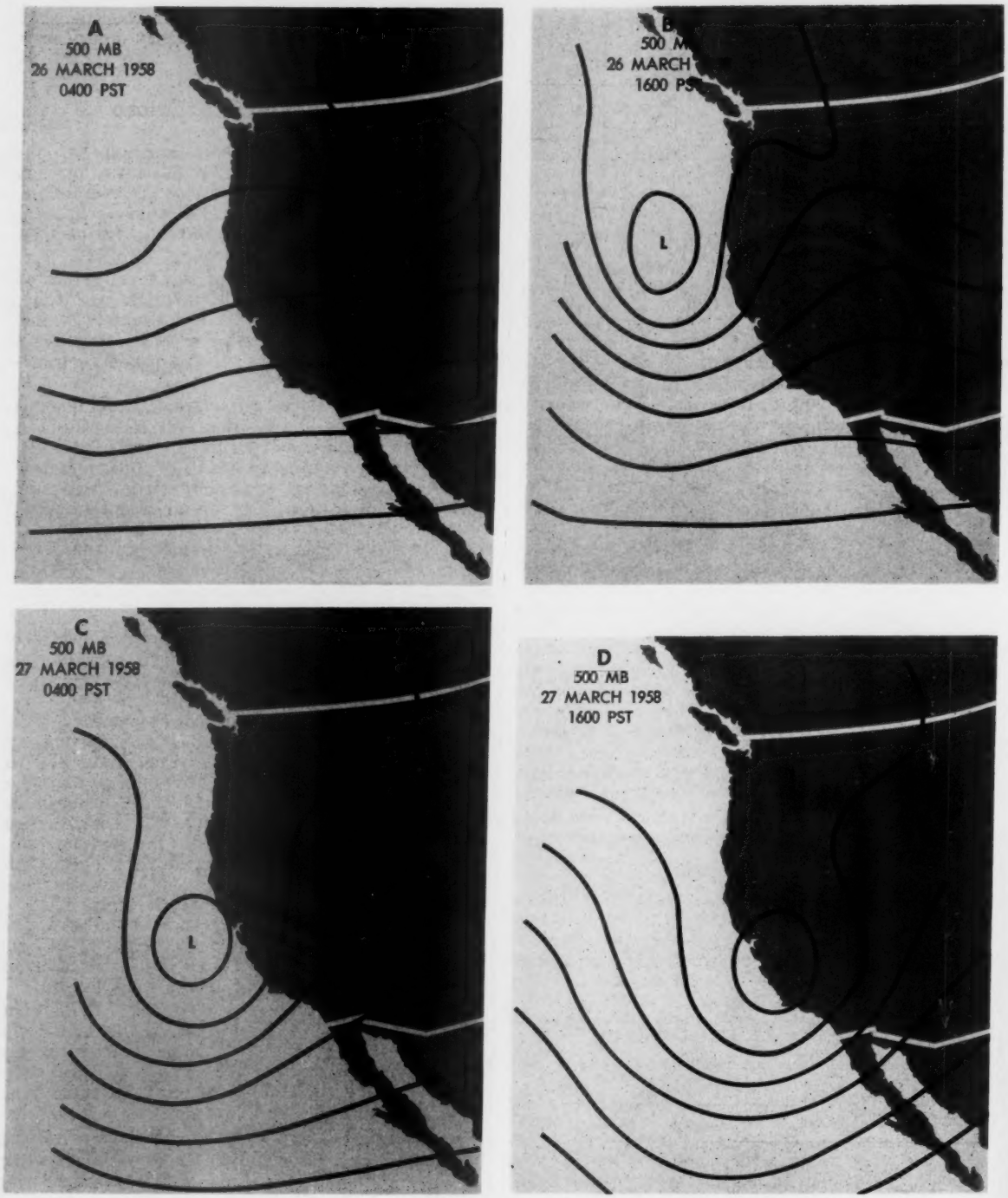


FIGURE 9. *Five hundred millibar pressure-surface contour charts at 12-hr intervals from 0400 P.s.t. 26 March to 0400 P.s.t. 28 March.*

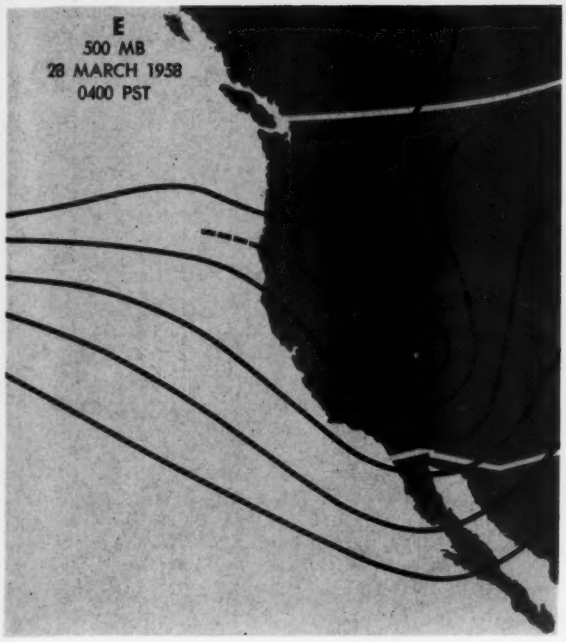


FIGURE 9.—Continued.

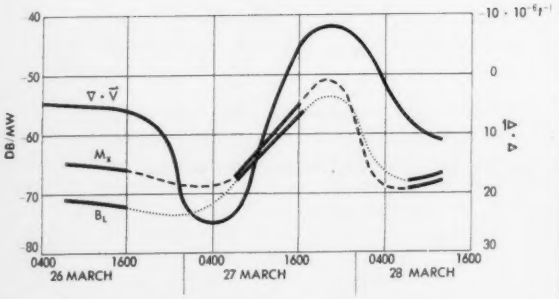


FIGURE 10. *Santa Barbara L-band and Point Mugu X-band signal levels and 500-mb velocity divergence for 0400 P.s.t. for 26 March to 1600 P.s.t. 28 March inclusive.*

Dashed signal level curves are extrapolated values for periods of radio silence.

6. Appendix A: Scattering Angle and Height of the Common Scattering Volume

The scattering volume is the portion of atmosphere common to both transmitting and receiving antenna beams. The acute angle at the beam intersection is called the scattering angle. It is possible to compute the height of the base of the scattering volume and the magnitude of the scattering angle from the gross refractive index profile reduced from radiosonde or refractometer data.

The atmospheric model (fig. 11) assumes horizontal homogeneity of a refractive layer for the distance that a ray segment remains within the layer. The number of refractive layers may vary with distance over the scattering link. The horizontal distribution of vertical refractive index gradients may be interpolated from soundings taken near either end of the radio link.

In figure 11, unprimed quantities refer to parameters between the transmitting (receiving) antenna and the beam intersection; primed quantities refer to parameters between the intersection and the receiving (transmitting) antenna. Symbol meanings are as follows:

A=antenna heights,
d_o=distance to the radio horizon,
d_i=projection on the earth of the ray path in the i-th layer,
D=distance between antennas,
w_i=thickness of i-th layer,
M=modified refractive index=$(n-1)\times 10^6 + 0.048\ h$,
h_v=height of base of the common volume,
θ=scattering angle.

Subscript v refers to the refractive layer wherein the beams intersect. From figure 11:

$$S=d_v+d_v'=D-\sum_0^{v-1} d_i. \qquad \text{(A1)}$$

The angle which the ray makes with the local horizontal upon emerging from the n-th layer is given by

$$d_n=10^{-3}\left[10^{+6}\alpha_0+2\sum_1^n W_i\frac{dM}{dh}\,i\right]^{\frac{1}{2}}. \qquad \text{(A2)}$$

For the case of grazing propagation α_o is the angle the ray makes with the horizon and is zero. For nongrazing propagation α_o is the elevation angle of the antenna beam. Figure 12 is a graph of $2WdM/dh$ as a function of w and $\Delta M/1{,}000$ ft. Figure 13 is a nomographic solution of (A2). The incremental distances d_o and d_1 can be easily determined from figure 14, the solution of

$$d=10^3\left[2H\left(\frac{dM}{dh}\right)^{-1}\right]^{\frac{1}{2}}. \qquad \text{(A3)}$$

The parameter H is either antenna height or the depth of the surface-based refractive layer at each end of the link.

The incremental distances d_i from d_2 to d_{v-1} may be determined from figure 15, a graphical solution of

$$d_i=10^6(\alpha_i-\alpha_{i-1})\left(\frac{dM}{dh}\right)^{-1}. \qquad \text{(A4)}$$

S may now be determined according to (A1). The remaining distances, d_v and d_v', are given by

$$d_\tau=\frac{\frac{S^2}{4}\left(\frac{dM}{dh}\right)_v 10^{-6}+S\alpha_{v-1}+\Delta W}{\frac{S}{2}\left(\frac{dM}{dh}\right)_v 10^{-6}+\alpha_{v-1}+\alpha'_{v-1}} \tag{A5}$$

and

$$d'_v=S-d_v. \tag{A6}$$

The height of the base of the common volume is now determined from

$$h_v=d_\tau\left(\alpha_{v-1}+10^6\frac{d_\tau}{2}\frac{dM}{dh}_v\right)+\sum_1^{v-1}W_i \tag{A7}$$

and the scattering angle from

$$\theta=\alpha_v+\alpha'_v+10^{-6}S\frac{dM}{dh}_v. \tag{A8}$$

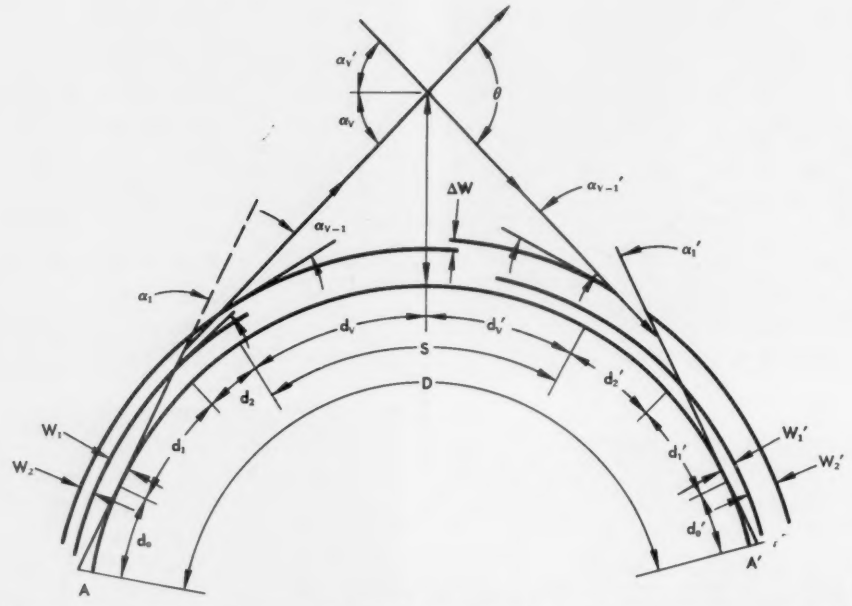


FIGURE 11. *Schematic diagram of scatter propagation link with nonhomogeneous, layered atmosphere.*

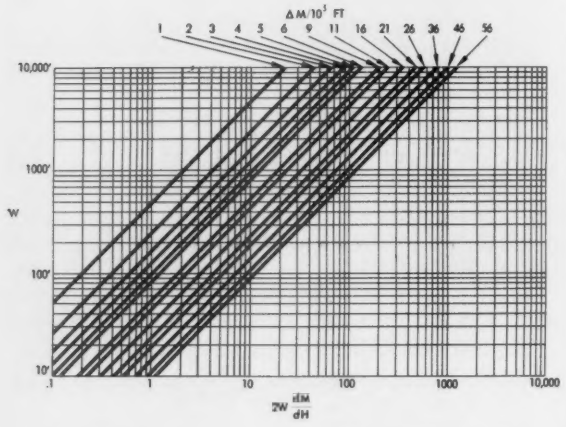


FIGURE 12. *Nomogram of* $2w\frac{dm}{dh}=f(w, \Delta M/10^3 \text{ ft})$.

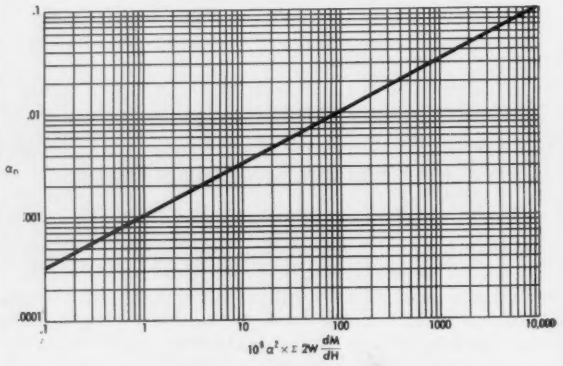


FIGURE 13. *Nomogram of* $\alpha_n=f(W_i, \Delta M/10^3 \text{ ft}, \alpha_\omega)$.

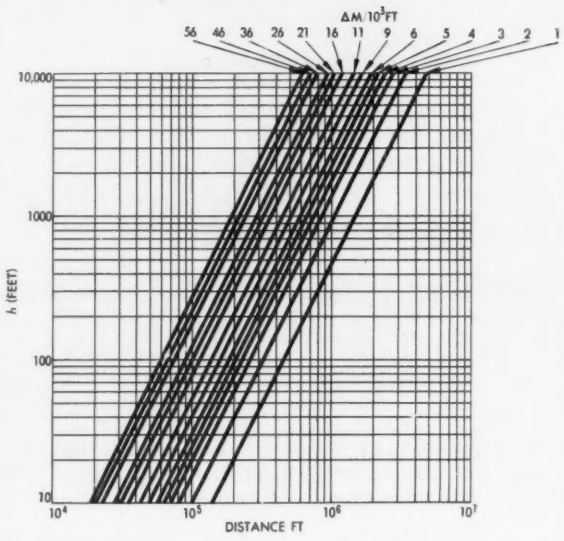


FIGURE 14. *Nomogram of* d=f(H, ΔM/10³ ft).

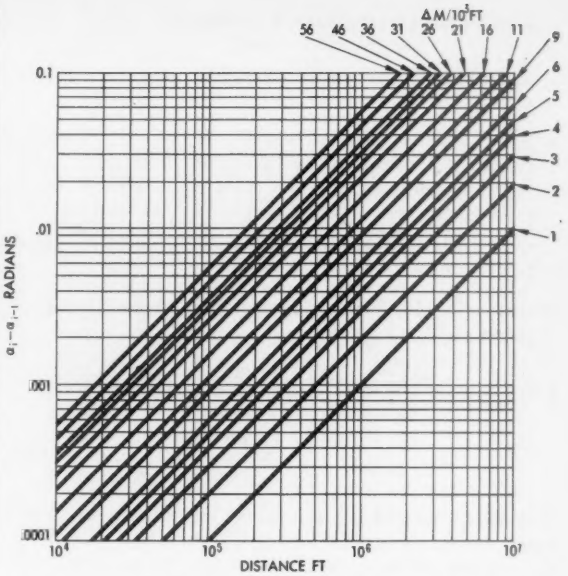


FIGURE 15. *Nomogram of* $d_i=(\alpha_i-\alpha_{i-1})(10^3\Delta M^{-1})$.

7. Appendix B

7.1. Method for Estimating the Relative Magnitude and Direction of the Vertical Component of the Wind

It may become necessary for the radio meteorologist to prepare operational propagation forecasts for areas where machine-computed vertical velocity charts are not available. The following simple method for obtaining a "better-than-guess" estimate of the vertical-motion field is suggested as a forecast aid.

The atmospheric model assumes the following conditions:

1. An equivalent barotropic level (or level of nondivergence) exists in the midtroposphere, usually at a lower altitude than the 500-mb surface and certainly below the 300-mb surface.
2. The horizontal wind velocity divergence changes sign at the level of nondivergence (from here on "LND") on the vertical coordinate and at no other elevation.
3. The direction and magnitude of vertical velocity throughout the troposphere are the same in sign and proportional in magnitude to those at the LND.
4. Vertical velocities are nonexistent at the surface of the earth and the outer boundary of the atmosphere.

One form of the equation of continuity is

$$\frac{\partial\rho}{\partial t}+\boldsymbol{V}\cdot\nabla\rho+\rho\nabla\cdot\boldsymbol{V}+\frac{\partial}{\partial Z}(\rho W)=0 \qquad \text{(B1)}$$

where ρ=density of the air, $\boldsymbol{V}$=horizontal wind velocity vector, ∇=horizontal gradient operator, Z=vertical coordinate, W=vertical wind velocity.

Integration of (B1) from the LND to the outer atmosphere results in

$$(\rho W)_{\text{LND}}=\int_{\text{LND}}^{\infty}\boldsymbol{V}\cdot\nabla\rho dZ+\int_{\text{LND}}^{\infty}\frac{\partial\rho}{\partial t}dZ+\int_{\text{LND}}^{\infty}\rho\nabla\cdot\boldsymbol{V}dZ. \qquad \text{(B2)}$$

It is usually observed that $\partial\rho/\partial t\ll\rho\nabla\cdot\boldsymbol{V}$ so that under most circumstances (B2-b) may be disregarded. It is also observed that cold air advection below the tropopause is accompanied by warm air advection in the stratosphere. For such a case (B2-a)$\rightarrow$0. With these assumptions (B2) reduces to

$$(W)_{\text{LND}}\cong\frac{1}{\rho}\int_{\text{LND}}^{\infty}\rho\nabla\cdot\boldsymbol{V}dZ. \qquad \text{(B3)}$$

Horizontal velocity divergence is defined as

$$\nabla\cdot\boldsymbol{V}=\frac{\partial v}{\partial s}+v\frac{\partial\theta}{\partial n}$$

in the natural coordinate system. Evaluation directly from the wind field is difficult to accomplish with any degree of accuracy. For this reason an alternate expression for the integrand of (B3) is sought.

The vorticity equation is expressed

$$\zeta^{-1}\frac{d\zeta}{dt}=-\nabla\cdot\mathbf{V} \tag{B4}$$

where ζ is the vertical component of the absolute vorticity. The substantial derivative equated to its partial components is

$$\frac{d\zeta}{dt}=\frac{\partial\zeta}{\partial t}+V\frac{\partial\zeta}{\partial s}+W\frac{\partial\zeta}{\partial Z} \tag{B5}$$

where s is along the streamlines. Substituting eq (B5) in (B3) we have

$$(W)_{\text{LND}}=-\frac{1}{\rho}\left[\int_{\text{LND}}^{\infty}\frac{\rho}{\zeta}\frac{\partial\zeta}{\partial t}dZ+\int_{\text{LND}}^{\infty}\frac{\rho}{\zeta}v\frac{\partial\zeta}{\partial s}dZ+\int_{\text{LND}}^{\infty}\frac{\rho}{\zeta}W\frac{\partial\zeta}{\partial Z}dZ\right]. \tag{B6}$$

The first integral on the right becomes proportionately smaller with increasing height above the LND and at the 300-mb level is generally quite small compared with the second integral. This is merely a statement that, in general, the winds blow through the pressure contours faster than the pressure contours are moving. The third integral is small because in the integral over the entire upper atmosphere the only contribution can pass through the lower boundary. Because the vertical velocity is several orders of magnitude smaller than horizontal components, the contribution cannot be great.

We can then write

$$(W)_{\text{LND}}\cong-\frac{1}{\rho}\int_{\text{LND}}^{\infty}\frac{\rho}{\zeta}V\frac{\partial\zeta}{\partial s}dZ. \tag{B7}$$

The absolute vorticity, $\zeta=f+\zeta_r$, and its gradient $\partial\zeta/\partial s=\partial f/\partial s+\partial\zeta_r/\partial s$ may be evaluated directly from a constant pressure chart. The gradient of the Coriolis term, $\partial\zeta/\partial s$ is, under most circumstances, at least an order of magnitude less than the gradient of relative vorticity, $\partial\zeta_r/\partial s$.

The magnitude of the vertical component of relative vorticity is $\zeta_r=(\partial v/\partial x)-(\partial u/\partial y)$ Substituting in eq (B7) we have finally

$$(W)_{\text{LND}}=-\frac{1}{\rho}\int_{\text{LND}}^{\infty}V\frac{\rho}{\zeta}\frac{\partial}{\partial s}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)dZ \tag{B8}$$

If the assumptions stated at first are reasonable then the sign of

$$(W)_{\text{LND}}=-\text{sign of}\left[\frac{\partial}{\partial s}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)\right]_h \qquad h>h_{\text{LND}}$$

Further, the 300-mb level is near the jetstream (region of maximum winds) and (with somewhat less certainty) one can also say

relative magnitude (W_{LND}=relative magnitude of

$$\frac{V}{\zeta}\frac{\partial}{\partial s}\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)\text{ 300 mb.}$$

7.2. Computation of ζ_r

The computation of relative vorticity on a constant pressure chart is greatly simplified if one makes use of the geostrophic wind approximation

$$U_g=-\frac{g}{f}\frac{dh}{dy}, \qquad V_g=\frac{g}{f}\frac{dh}{dx}.$$

The finite difference equation for the relative vorticity is then

$$\zeta_r=\frac{g}{f}\left(\frac{\Delta h}{(\Delta x)^2}+\frac{\Delta h}{(\Delta y)^2}\right) \tag{B9}$$

The vorticity field is computed by reading the constant pressure surface height at evenly spaced grid points and carrying out the computation

$$\zeta_r=\frac{g}{fD^2}\left(H_1+H_2+H_3+H_4-4H_5\right)$$

where f is the Coriolis parameter (twice the angular velocity of the earth times the sine of the latitude), g the acceleration of gravity, D is the grid spacing, and H_1, H_2, H_3, H_4, H_5 are the pressure surface heights at the points P_1 . . . P_5 as shown in figure 16. The computed vorticity is the average value for the grid space and is plotted at P_5.

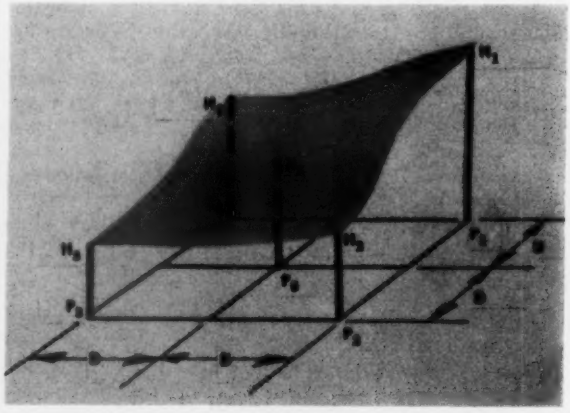


FIGURE 16. *Schematic of grid used to compute relative vorticity on a constant pressure surface.*

In this report the relative vorticity was computed using a grid size of 5° latitude for a grid centered over Catalina Island and another grid 5° latitude upwind. The vorticity gradient was assumed to be the finite difference between ζ_r at the two positions divided by the 5° latitude distance between points.

8. References

[1] B. R. Bean, Some meteorological effects on scattered radio waves, IRE Trans. CS–**4,** 32 (1956).

[2] J. B. Smyth and others, Atmospheric refraction; a preliminary quantitative investigation, Navy Radio and Sound Lab. Rept. WP–17 (28 Dec. 1944).

[3] J. B. Smyth and L. G. Trolese, Propagation of radio waves in the lower troposphere, Proc. I.R.E. **35,** 1198 (1947).

[4] E. E. Gossard and L. J. Anderson, The effect of superrefractive layers on 50–5,000 Mc non-optical fields, IRE Trans. AP–**4,** 175 (1956).

[5] H. G. Booker and W. E. Gordon, A theory of radio scattering in the troposphere, Proc. I.R.E. **38,** 401 (1950).

[6] F. Villars and V. F. Weisskopf, The scattering of electromagnetic waves by turbulent atmospheric fluctuations, Phys. Rev. **94,** 232 (1954).

[7] G. K. Batchelor, The scattering of radio waves in the atmosphere by turbulent fluctuations in refractive index, Cornell Univ., School of Elec. Eng., Tech. Rept. 26 (15 Sept. 1955).

[8] J. R. Bauer, The suggested role of stratified elevated layers in transhorizon short-wave radio propagation, Mass. Inst. Technol., Lincoln Lab., Tech. Rept. 124 (24 Sept. 1956).

[9] H. T. Friis and others, A reflection theory for propagation beyond the horizon, Bell System Tech. J. **36,** 627 (May 1957).

[10] D. L. Ringwalt and others, Measurements of Mc/s scatter propagation as function of meteorology (letter submitted to IRE for PGAP series).

[11] P. Misme, The correlation between the electric field at a great distance and a new radio-meteorological parameter, IRE Trans. AP–**6,** 289 (1958).

[12] W. F. Moler and W. A. Arvola, Vertical motion in the atmosphere and its effect on VHF radio signal strength, Trans. Am. Geophys. Union **37,** 399 (1956).

[13] E. E. Gossard, Power spectra of temperature, humidity, and refractive index from aircraft and tethered balloon measurements (to be published).

[14] W. E. Gordon, Radio scattering in the troposphere, Proc. I.R.E. **43,** 23 (1955).

[15] J. Michaelis and E. E. Gossard, Distribution of refractive layers over the North Pacific and Arctic, Navy Electron. Lab. Report 841 (18 Apr. 1958).

BOULDER, COLO.

(Paper 64D1–42)

Journal of Research of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 1, January–February 1960

Layered Earth Propagation in the Vicinity of Point Barrow, Alaska[1]

Glenn M. Stanley[2]

(September 1, 1959)

The relative field strength of a vertically polarized low-frequency radio signal was measured as a function of distance over several radial paths in the vicinity of Point Barrow, Alaska. The attenuation of the recorded signal was very much less than predicted by the theory of propagation of a ground wave signal over a plane, homogeneous, infinitely conducting earth. The analysis of these data in terms of a plane, layered, finitely conducting earth appears to resolve the anomaly.

1. Introduction

For several years, studies of effective ground conductivities in Alaska have been in progress by the Geophysical Institute at the University of Alaska. The project has been carried on under the auspices of the U.S. Navy and the Arctic Institute of North America. The data used in this paper were gathered under a Buships contract administered by the Signal Corps.

The relative field strength, as a function of distance from the transmitting antenna, was measured along several radials from each of 25 Civil Aeronautics Administration radio ranges and beacons in Alaska. These ranges and beacons operate in the frequency range of from 200 to 400 kc. The measurements were made from a low-flying aircraft equipped with a sensitive receiver, an Esterline-Angus recorder, and a nearly vertical trailing wire antenna. The effective ground conductivity is obtained by a comparison of the experimental relative field strength values with families of theoretical curves obtained by Norton's method. From this comparison, the value of conductivity that best fits the experimental data is taken as the value of effective conductivity over the path.

Signals were recorded continuously along radii from the transmitting antennas for distances up to 120 miles. Check points along the flight paths were obtained from U.S.G.S. 1:250,000 topographical contour maps. The aircraft used was equipped with a gyrocompass, and constant course checks were made with maps. In most instances, deviations from the radial course were less than a few hundred feet.

As all signals were recorded along radii from the transmitting antenna, no corrections were necessary for assymetrical transmitting and receiving antennas. At a distance of approximately 30 wavelengths, no appreciable change in the signal level was apparent in a ½ mile horizontal deviation from the radial flight path. In addition, no appreciable changes in the signal levels were noted for changes of a few hundred feet in elevation at the same distance. The elevation of the receiving antenna, however, was kept to within 400 ft of the local terrain (and in general, was nearer to 200 ft) to reduce the effect of interference between the direct and the reflected signal.

Relative field strengths were used throughout the computations of effective conductivity and no attempt was made to measure absolute field strength.

Measurements made in the vicinity of Point Barrow, Alaska, showed that the signals did not attenuate as rapidly as might have been expected for a region known to be underlain with permafrost to depths of 500 to 1,000 ft. Flight runs were made from Point Barrow in June and again in August. With one exception, the signal strength decreased less rapidly than predicted by the theory of a plane homogeneous earth for an infinitely conducting ground.

2. Theoretical Consideration

To make clear the implications of the above mentioned anomaly, it is appropriate to review briefly the theory underlying the experimental measurements. The surface wave field strength at short distances is given by the relation

$$E=2E_o|F_e| \text{ and } E_o=\frac{\text{const.}}{\rho}, \tag{1}$$

where E_o=the maximum field strength of the antenna if it were located in free space
ρ=the distance from the transmitting antenna to the receiving antenna and
$|F_e|$=the surface wave attenuation factor;

provided the antennas are very near the ground and that

$$2\lambda<\rho<\frac{50}{f^{1/3}\text{ mc}}\text{ mi}. \tag{2}$$

[1] This paper was presented at the Symposium on VLF Radio Waves held in Boulder, Colo., January 1957 (invited paper).
[2] Present address: Arctic Aeromedical Laboratory, APO 731, Seattle, Wash.

The surface wave attenuation factor may be obtained by methods given by Norton [1].[3]

Taking the ratio of the relative field strengths at two points at distance ρ_n and ρ_x respectively, from the transmitting antenna, where ρ_n is greater than ρ_x, it may be seen that

$$\frac{E_n}{E_x}=\frac{|F_e|_n\rho_x}{|F_e|_x\rho_n}. \tag{3}$$

Using this relation a relative field strength curve may be drawn from any convenient initial point. Each set of values of ground constants will give one such curve. Figure 1 shows a family of these curves for $f=236$ kc, $\epsilon=5\ \epsilon_o$ and an initial point of 2.3 mi from the transmitting antenna.

The surface wave attenuation factor F_e is given by

$$F_e=1-i(\pi\ p_e)^{1/2}e^{-p}\ \text{erfc}\ (ip_e^{1/2}), \tag{4}$$

where p_e is the effective numerical distance and is given by

$$p_e=|p_e|e^{ib}. \tag{5}$$

Assuming a homogeneous earth, b is negative and $|F_e|$ is less than unity. However, Wait has shown that (for certain conditions of horizontally stratified media) b may be positive and $|F_e|$ greater than unity [2, 3, 4]. Wait presents tables of $|F_e|$ as a function of $|p_e|$ for positive values of b to $+75$ degrees and for negative values of b to -90 degrees. The negative values were obtained using the results of Norton. $|p_e|$ can be computed in terms of the impedance of the surface of the top layer looking normally downward. The impedance of the top layer is related to the impedance of each succeeding lower layer by a recurrence equation. For a two-layer earth, where $\sigma_2>>\sigma_1$, and h_1 is not too small,

$$|p_e|=\frac{\pi\rho\mu\omega}{(120\pi)^2\lambda_0\sigma_1}(A^2+B^2); \tag{6}$$

$$b=\tan^{-1}\left(\frac{2AB}{A^2-B^2}\right); \tag{7}$$

where

$$A=\frac{\tanh w\sec^2 w}{1+\tanh^2 w\tan^2 w}; \tag{8}$$

$$B=\frac{\tan w\,\text{sech}^2 w}{1+\tanh^2 w\tan^2 w}; \tag{9}$$

and

$$w=\frac{(\sigma_1\mu w)^{\frac{1}{2}}h_1}{\sqrt{2}}. \tag{10}$$

From these relations the effective numerical distance $|p_e|$, and the phase angle b, may be obtained for a given frequency, a given thickness and conductivity of the upper layer. When $|p_e|$ and b are known, the surface wave attenuation factor $|F_e|$ can be determined. Since $|F_e|$ is the only unknown entering the equation for the relative field strength, curves may then be drawn for different values of the parameters σ_1 and h_1. Figure 2 shows curves of relative field strength for some values of h_1 taking σ_1 to be constant and equal to 1 mmho/m.

3. Results and Discussion

Returning to the experimental values of relative field strength obtained near Point Barrow, an attempt is made to explain the low attenuation of the signals in terms of layered earth propagation.

The existence of highly conducting saline layers or lenses at depths of from 10 to 100 ft in the immediate vicinity of Point Barrow is deduced from test wells driven by the Arctic Contractors during the course of their work on U.S. Navy Petroleum Reserve #4. The extent and exact nature of these layers or lenses is not known. However, relative field strength measurements indicate that the propagation characteristics change markedly at distances of from 60 to 100 mi southward from Point Barrow. In addition, there is visual evidence of an abandoned beach line at about this same location. Mr. Max Brewer, of the Arctic Research Laboratory at Point Barrow reports that the salinity of these layers near Point Barrow is equal to, or greater than that of sea water.

Figure 3 shows flight runs 93–95 made in June 1954. The vertical lines depict the spread of the three runs while the circles represent average values. Other flight runs in the vicinity have shown similar conditions of anomalous propagation.

Thus, it is reasonable to approximate the ground in the vicinity of Point Barrow by a two layer ground. The conductivity of the lower layer has been assumed to be equal to that of sea water (5000 mmho/m). Several families of relative field strength curves were drawn for different values of the conductivity of the upper layer σ_1, and for different thicknesses of the upper layer h_1. The thickness of the lower layer was assumed to be semi-infinite. Figure 4 shows the average values obtained from flight runs 89–92 and 202–205 superimposed on the relative field strength curves for σ_1 equal to 10 mmho/m. The comparison of the experimental curve with the theoretical curves suggests that the apparent thickness of the upper layer h_1 is equal to approximately 9 m.

Flight runs 93–95 and 197–199, which were about 30 deg to the East, and flight runs 193–196, approximately 30 deg to the West of those shown in figure 2 have similar characteristics. In contrast, flight runs 85–88, show a very high signal level. No combination of ground constants that has been tried will produce a relative field strength curve which matches the experimental curve corresponding to this high signal level.

[3] Figures in brackets indicate the literature references at the end of this paper.

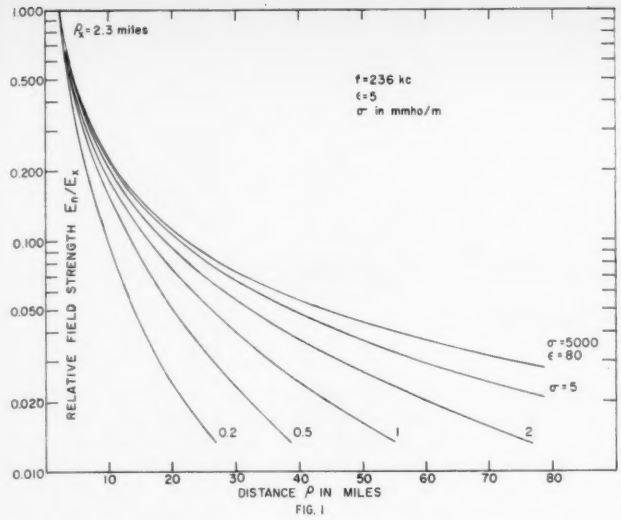


FIG. 1

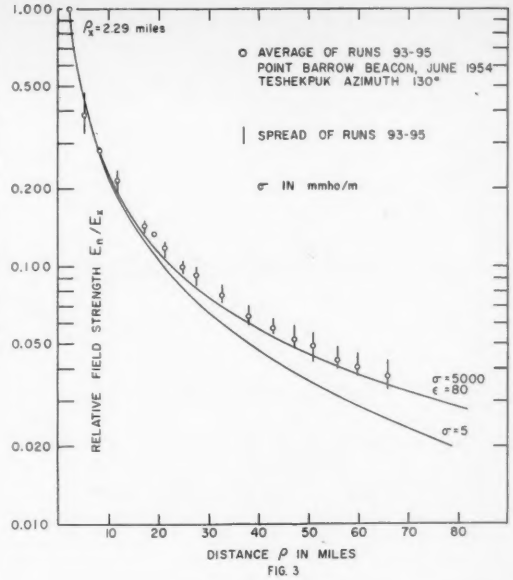


FIG. 3

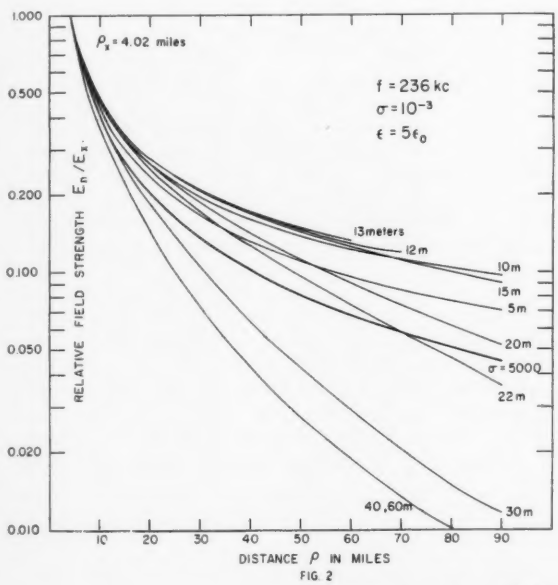


FIG. 2

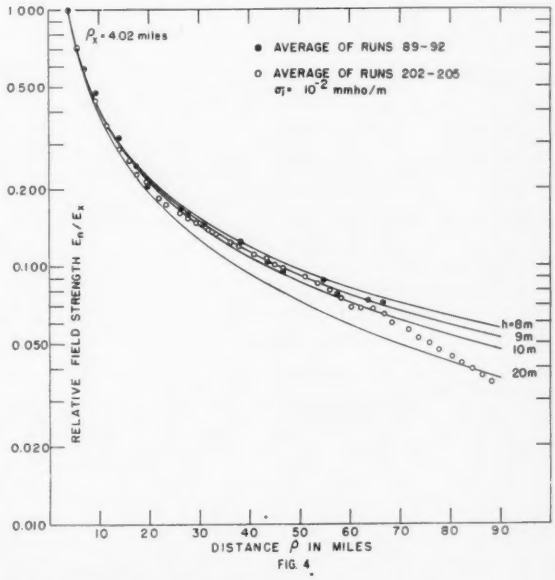


FIG. 4

FIGURES 1, 2, 3, 4. *Field strength versus distance curves.*

The analysis of these data indicates, that at a frequency of 236 kc, there is, indeed, layered earth propagation in the vicinity of Point Barrow, Alaska. The analysis further suggests that the ground conductivity of the upper 9 m is approximately 10 mmho/m, when measured by the method outlined.

The area of measurement which encompasses most of the Arctic slope is essentially flat and free from manmade interference such as fences, railroads, powerlines, and other sources of disturbance. It therefore appears that this area would provide an excellent test site for various low-frequency propagation studies.

4. References

[1] K. A. Norton, The calculation of ground-wave field intensity over a finitely conducting spherical earth, Proc. IRE, **29,** 623 (1941).

[2] James R. Wait, Radiation from a vertical electric dipole over a stratifield ground, Trans. IRE PGAP, **AP–1,** 9 (1953).

[3] James R. Wait and W. C. G. Fraser, Radiation from a vertical dipole over a stratified ground (part II), Trans. IRE PGAP, **AP–3,** 144 (1954).

[4] James R. Wait, The theory of electromagnetic surface waves of geological conductors, Geofis. pura e appl. (Milano), **28,** 47 (1954).

BOULDER, COLO.

(Paper 64D1–43)

Announcement of Systematic Ionospheric Electron Density Data

Reduction of hourly ionospheric vertical soundings to electron density profiles has become a part of the systematic ionospheric data program of the Central Radio Propagation Laboratory, National Bureau of Standards. Scalings of ionograms for this purpose are being provided by ionosphere stations operated by CRPL and the U.S. Army Signal Corps. For the present, the hourly profile data from one CRPL station, Puerto Rico, are appearing in the monthly CRPL-F Reports, Part A.[1] These data are in place of the standard ionogram reductions formerly provided by this station. The very considerable task of scaling the ionograms for this purpose is being undertaken by T. R. Gilliland, Engineer in Charge, Puerto Rico Ionosphere Sounding Station; the computations are performed at the NBS Boulder Laboratories by a group headed by J. W. Wright. Basic conversion of virtual to true heights uses the well-known matrix method developed by K. G. Budden of the Cavendish Laboratory, Cambridge University, programed for an IBM 650 computer.

The quantities given for each hour of each day of the month are: The electron density at each 10-km interval of height, including the maximum electron density proportional to $(foF2)^2$; h_{min}, the height of zero or very low electron density, obtained by linear extrapolation of the $N(h)$ curve; h_{max}, the height of maximum electron density, determined by fitting a parabola to the upper portion of the profile; and $S(h_{max})$, the integrated electron density between h_{min} and h_{max}.

Tabulations of two arithmetic mean electron densities are also given for each hour. One average is for the "undisturbed" ionosphere which includes the soundings taken when the magnetic character figure Kp is less than $4+$; the other includes the remaining data to form a "disturbed" average. The latter may have little physical significance because the number of disturbed hours is usually small and the behavior of the ionosphere during disturbed hours is not consistent.

Before the averaging process, the individual profiles are extrapolated above h_{max} by a Chapman distribution of 100-km scale height. This assumed model seems to agree well with the few published measurements dealing with the topside profile of the F-region. Extrapolation is necessary in order to calculate homogeneous averages near h_{max} and the average profiles are, in fact, given up to 950 km. Also given are the integrated electron densities estimated to infinity, $S(h_{inf})$; this is an approximation to the total electron content in a vertical column of the ionosphere of 1 cm² cross section.

[1] The CRPL-F Reports, Part A, Ionospheric Data, have a limited press run and normally are available only on an exchange basis. However, compilations of tables of medians of ionospheric data and of tabulations of electron density appearing in this publication may be purchased in booklet form at the price of $1.10 per booklet. Please make inquiry of: IGY World Data Center A, Airglow and Ionosphere, Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

Publications of the National Bureau of Standards

Selected Abstracts

Earth currents near a monopole antenna with symmetrical top loading, James R. Wait, *J. Research NBS* **62,** *247* (*June 1959*).

Expressions for the fields are developed for a vertical ground-based monopole with a cone or disk located at the top of the antenna to simulate umbrella top loading. The current distribution on the structure is assumed. Using spherical-wave functions, the magnetic-field distribution on the ground plane near the base of the antenna is computed and illustrated by graphs. For the case where the antenna is electrically small, the currents flowing on the cone or disk are shown to contribute only slightly to the total field.

Earth currents near a top-loaded monopole antenna with special regard to electrically small L- and T-antennas, H. Lottrup Knudsen, *J. Research NBS* **62,** *283* (*June 1959*).

An investigation has been made of the ground currents near a top-loaded monopole with nonazimuthal symmetry. Formulas have been developed for the surface current density produced by an inclined, straight wire over a horizontal ground plane for an arbitrary current distribution on the antenna. Working formulas have been developed and numerical calculations of the surface current density on the ground plane have been carried out for the case of a small antenna with a linear current distribution. These results have been used for the calculation of the contribution to the surface current density due to the top loading in the case of an L-antenna and in the case of a T-antenna. In each case both the absolute value of the surface current density arising from the top loading and the relative value of its ϕ-component have been plotted. This component under certain circumstances may be important in calculating the ground losses in the case of a system of radial ground wires.

World maps of F2 critical frequencies and maximum usable frequency factors, D. H. Zacharisen, *NBS 2 Tech. Note* (*PB151361*) *$3.50.*

This report was prepared for the purpose of presenting six months of contour maps and charts for use in predicting F2-layer maximum usable frequencies. Prediction maps for each even hour of Greenwich Mean Time and charts in which time is continuous along the abscissa are given for the months of January, March, June, July, September, and December. The four parameters used for predicting MUFs are foF2 and the 4000 km MUF factor for a twelve-month running average Zurich sunspot number of 50, and the rates of change of foF2 and 4000 km MUF factor with sunspot number. The first three parameters use a map presentation with GMT constant over the surface of the map. The fourth parameter uses a chart presentation in which the ordinate is geomagnetic latitude and the abscissa is local time.

Low- and very-low-radiofrequency tables of ground-wave parameters for the spherical earth theory: the roots of Riccati's differential equation (*supplementary numerical data for NBS Circ. 573*), J. R. Johler, L. C. Walters, C. M. Lilley, *NBS Tech. Note* (*PB151366*) *$2.25.*

The roots of Riccati's differential equation are tabulated in detail for the low- and very-low radiofrequencies. The zeroes and certain other parameters used in the calculation of the amplitude and phase of the ground wave by the conventional "residue series" are also tabulated.

Green coronal line intensity and geomagnetism, C. Warwick, *J. Geophys. Research* **64,** *527* (*May 1959*).

Climax Observatory spectrograms of the green coronal line, λ5303, for the period 1942 to 1944 near solar activity minimum have recently been remeasured. Analysis of these revised intensities indicates that the most consistent feature in the relation of green line intensity to geomagnetism is a minimum in geomagnetic activity following central meridian passage (CMP) of regions of high green line intensity. This effect may be interpreted as a result of the tendency of solar activity regions to coincide with regions of high green line intensity and to avoid M regions. In the period October 1953 to October 1954, at sunspot minimum, no relation appeared between green line intensity and geomagnetic activity.

Calculated patterns of slotted elliptic-cylinder antennae, James R. Wait and Walter E. Mientka, *Appl. sci. Research, Sec. B* **7,** *449* (*1959*).

The model assumed is a perfectly conducting elliptic cylinder of infinite length which has a narrow axial slot of finite length. Patterns are presented for various ratios of the major to minor axis. It is indicated that for reasonably large elliptic cylinders the patterns are dependent mainly on the surface curvature in the neighborhood of the slot.

A study of limb flares and associated events, C. Warwick and M. Wood, *Astrophys. J.* **129,** *801* (*May 1959*).

Measured heights of limb flares were used to derive a frequency distribution of flare height and to assess the importance of height in the production of flare-associated events. The observed height distributions can be explained by a real distribution with maximum at zero, decreasing toward greater heights. The occurrence of SWF (short-wave fadeout) depends strongly on flare height. This dependence could be explained as an effect either of absorption of the ionizing radiation in the sun's atmosphere or of a mechanism of production of ionizing radiation that depends critically on the height of the flare. Center-to-limb variation of SWF occurrence, while not conclusive, favors the former interpretation. The relation between optical flare characteristics and occurrence of associated bursts of radionoise is strongest for decimeter bursts and weaker for bursts at meter wave lengths.

Some characteristics of VLF propagation using atmospheric waveforms, William L. Taylor and L. Jerome Lange, *Recent advances in atmospheric electricity, p. 609* (*Pergamon Press, 1959*).

Simultaneous observations have been made of the waveforms of atmospherics ("sferics") at four widely separated stations. VLF attenuation characteristics have been computed by comparing the spectra of the waveforms recorded at the four locations for each discharge. The results of some preliminary measurements of VLF attenuation are presented for the band of frequencies from 4 kc to 30 kc at distances of 1,200 km to 6,500 km.

Other NBS Publications*

Journal of Research, Section 63A, No. 3, November–December 1959. 70 cents.

Multiple ionization of rare gases by electron impact. M. Krauss, R. M. Reese, and V. H. Dibeler.

Light scattering by commercial sugar solutions. Carl J. Rieger and Frank G. Carpenter.

Analysis of the first spectrum of ruthenium (Ru I). K. G. Kessler.

Supplementary Zeeman data for the first spectrum of ruthenium (Ru I). J. Rand McNally, Jr., and Karl G. Kessler.

Low even configurations in the first spectrum of ruthenium (Ru I), part 2. R. E. Trees.

Thermal degradation of polymers at high temperatures. Samuel L. Madorsky and Sidney Straus.

Influence of impurities on the pyrolysis of polyamides. Sidney Straus and Leo A. Wall.

A preliminary list of levels and *g*-values for the first spectrum of thorium (Th I). Romuald Zalubas.

OH in the solar spectrum. Charlotte E. Moore and Herbert P. Broida.

Use of Chebychev polynomials in thin film computations. Klaus D. Mielenz.

World maps of F2 critical frequencies and maximum usable frequency factors, D. H. Zacharisen, NBS Tech. Note 2 (PB151361) $3.50 (see abstracts above).

Low- and very low-radiofrequency tables of ground wave parameters for the spherical earth theory: The roots of Riccati's differential equation. (Supplementary numerical data for NBS Circ. 573) J. R. Johler, L. C. Walters, and C. M. Lilley, NBS Tech. Note 7 (PB151366) $2.25 (see abstracts above).

Calculated behavior of a fast neutron spectrometer based on the total absorption principle, J. E. Leiss, NBS Tech. Note 10 (PB151369). $1.00.

Penetration of gamma rays from isotropic sources through aluminum and concrete, M. J. Berger and L. V. Spencer, NBS Tech. Note 11 (PB151370). 50 cents.

Prediction of the cumulative distribution with time of ground wave and tropospheric wave transmission loss. Part 1. The prediction formula, P. L. Rice, A. G. Longley, and K. A. Norton, NBS Tech. Note 15 (PB151374). $1.50.

The changing character of chemical research in government, E. Wichers, The Chemist XXXVI, No. 7, 260 (1959).

Study of the setting of plaster, K. D. Jorgensen and A. S. Posner, J. Dental Research **38,** No. 3, 491 (1959).

Radiation attenuation data, H. O. Wyckoff, Radiation Hygiene Handbook, edited by H. Blatz, Section 8, p. 1 (McGraw-Hill Book Co., New York, N.Y., 1959).

Corrosion of type 310 stainless steel by synthetic fuel oil ash, H. L. Logan, Corrosion **15,** 443t (1959).

Physical research, Part 2, G. C. Paffenbarger and W. Souder, J. Am. Dental Assoc. **58,** 98 (1959).

Effects of gamma radiation on collagen, J. Cassel, J. Am. Leather Chemists' Assoc. LIV, No. 8, 432 (1959).

Standards for neutron flux measurement and neutron dosimetry, R. S. Caswell, E. R. Mosburg, Jr., and J. Chin, 2d United Nations International Conf. on the Peaceful Uses of Atomic Energy, Vol. 21, Health and Safety: Dosimetry and Standards, P/752 USA, p. 92 (1959).

Note on bivariate linear interpolation for analytic functions, W. Gautschi, Math. Tables and Other Aids to Computation **13,** No. 66, 91 (1959).

Absorption of radiation by a cylindrical sample of a strong absorber, P. H. Fang and I. A. Stegun, Letter to Editor, J. Chem. Phys. **31,** 267 (1959).

PILOT—A new multiple computer system, A. L. Leiner, W. A. Notz, J. L. Smith, and A. Weinberger, J. Assoc. Computing Mach. **6,** No. 3, 313 (1959).

Properties of rutile (titanium dioxide), F. A. Grant, Rev. Modern Phys. **31,** 646 (1959).

Analysis of vibrational relaxation data in shock wave experiments, K. E. Shuler, J. Chem. Phys. **30,** No. 6, 1631 (1959).

The interlaboratory evaluation of testing methods, J. Mandel and T. W. Lashof, ASTM Bull. No. 239, 53 (1959).

Branched-chain higher sugars. I. A 9-aldo-4-C-formyl-nonose derivative, R. Schaffer and H. S. Isbell, J. Am. Chem. Soc. **81,** 2178 (1959).

The structure of electrolytic solutions, W. J. Hamer (John Wiley & Sons, Inc., New York, N.Y., 1959).

The use of a vacuum microbalance in studies of electron tube materials, G. F. Rouse, Proc. 4th Natl. Conf. on Tube Techniques, September 10–12, 1958, Sponsored by the Advisory Group on Electron Tubes, p. 262 (New York University Press, New York, N.Y., 1959).

Widths of cracks in concrete at the surface of reinforcing steel evaluated by means of tensile bond specimens, D. Watstein and R. G. Mathey, J. Am. Concrete Inst. **31,** No. 1, 47 (1959).

Exposure standards and radiation protection regulations, L. S. Taylor, Radiation Hygiene Handbook, edited by H. Blatz, Section 3, p. 2 (McGraw-Hill Book Co., New York, N.Y., 1959).

Geometrical anisotropy of magnetic materials in wave guides and cavities, A. L. Steinert, J. Appl. Phys. **30,** 1109 (1959).

Radioactivity standardization in the United States, W. B. Mann and H. H. Seliger, 2d U.N. Intern. Conf. on the Peaceful Uses of Atomic Energy, Vol. 21, Health and Safety: Dosimetry and Standards, P/750 USA, p. 90 (1959).

Magnetic interaction of H_3, V. Griffing, J. L. Jackson, and B. J. Ransil, J. Chem. Phys. **30,** No. 4, 1066 (1959).

Program of the International Commission on radiological units and measurements, L. S. Taylor, L. H. Gray, and H. O. Wyckoff, 2d U.N. Intern.

Conf. on the Peaceful Uses of Atomic Energy, Vol. 21, Health and Safety: Dosimetry and Standards, P/2243 WHO, p. 81 (1959).

New approach in the theory of satellite orbits, J. P. Vinti, Phys. Rev. Letters **3,** No. 1, S517 (1959).

The nature, cause and effect of the porosity in electro-deposits, IV. Influence of gas bubbles on the formation of pores, F. Ogburn and D. W. Ernst, Plating **46,** 957 (1959).

Low even configurations in the first spectrum of ruthenium (Ru I), R. E. Trees, J. Opt. Soc. Am. **49,** 838 (1959).

End plate modification of X-band TE_{011} cavity resonators, M. C. Thompson, Jr., F. E. Freethey, and D. M. Waters, Letter to editor, IRE Trans. on Microwave Theory Tech. **MTT–7,** 388 (July 1959).

**Publications for which a price is indicated (except for Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (order by PB number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

Notice

The papers listed below will appear in early issues of the Journal of Research, Section D, Radio Propagation:

A theory of wavelength dependence in UHF transhorizon propagation based on meteorological considerations. Ralph Bolgiano, Jr.

Measured statistical characteristics and narrow-band teletype message errors on a single-sideband 600-mile-long UHF tropospheric link. E. F. Florman and R. W. Plush.

The effect of receiver bandwidth on amplitude distribution of VLF atmospheric noise. Forrest F. Fulton, Jr.

The electric field at the ground plane near a top-loaded monopole antenna with special regard to electrically small L- and T-antennas. H. Lottrup Knudsen and T. Larsen.

The joint use of the ordinary and extraordinary virtual height curves in determining ionospheric layer profiles. L. R. O. Storey.

Ionospheric effects due to nuclear explosions. W. F. Utlaut.

Impedance of a corner-reflector antenna as a function of the diameter and length of the driven element. A. C. Wilson.

Seventy-five kilocycles per second pulsed normal incidence ionospheric sounding. A. J. Ferraro, J. J. Gibbons, R. E. Houston, W. J. Ross, and E. R. Schmerling.

Preliminary radio meteorological study. Fumio Ikegami.

On the theory of reflection of low- and very-low radiofrequency waves from the ionosphere. J. Ralph Johler and Lillie C. Walters.

The tradewind inversion as a transoceanic duct. M. Katzin, H. Pezzner, B. Y.-C. Koo, J. V. Larson, and J. C. Katzin.

The aurora of October 22/23, 1958, at Rapid City, South Dakota. F. E. Roach and E. Marovich.

Focusing, defocusing, and refraction in a circularly stratified atmosphere. K. Toman.

Terrestrial propagation of VLF radio waves—a theoretical investigation. James R. Wait.

Relation of turbulence theory to ionospheric scatter propagation experiments. Albert D. Wheelon.

Impedance characteristics of a uniform current loop having a spherical core. Saburo Adachi.

On the calculation of the departures of radio wave bending from normal. B. R. Bean and E. J. Dutton.

Response of a loaded electric dipole in an imperfectly conducting cylinder of finite length. Charles W. Harrison, Jr. and Ronold W. P. King.

Optimum frequencies for outer-space communication. George W. Haydon.

An analysis of propagation measurements made at 418 Mc well beyond the radio horizon (a digest). H. B. Janes, J. C. Stroud, and M. T. Decker.

UNITED STATES
GOVERNMENT PRINTING OFFICE
DIVISION OF PUBLIC DOCUMENTS
WASHINGTON 25, D.C.

OFFICIAL BUSINESS

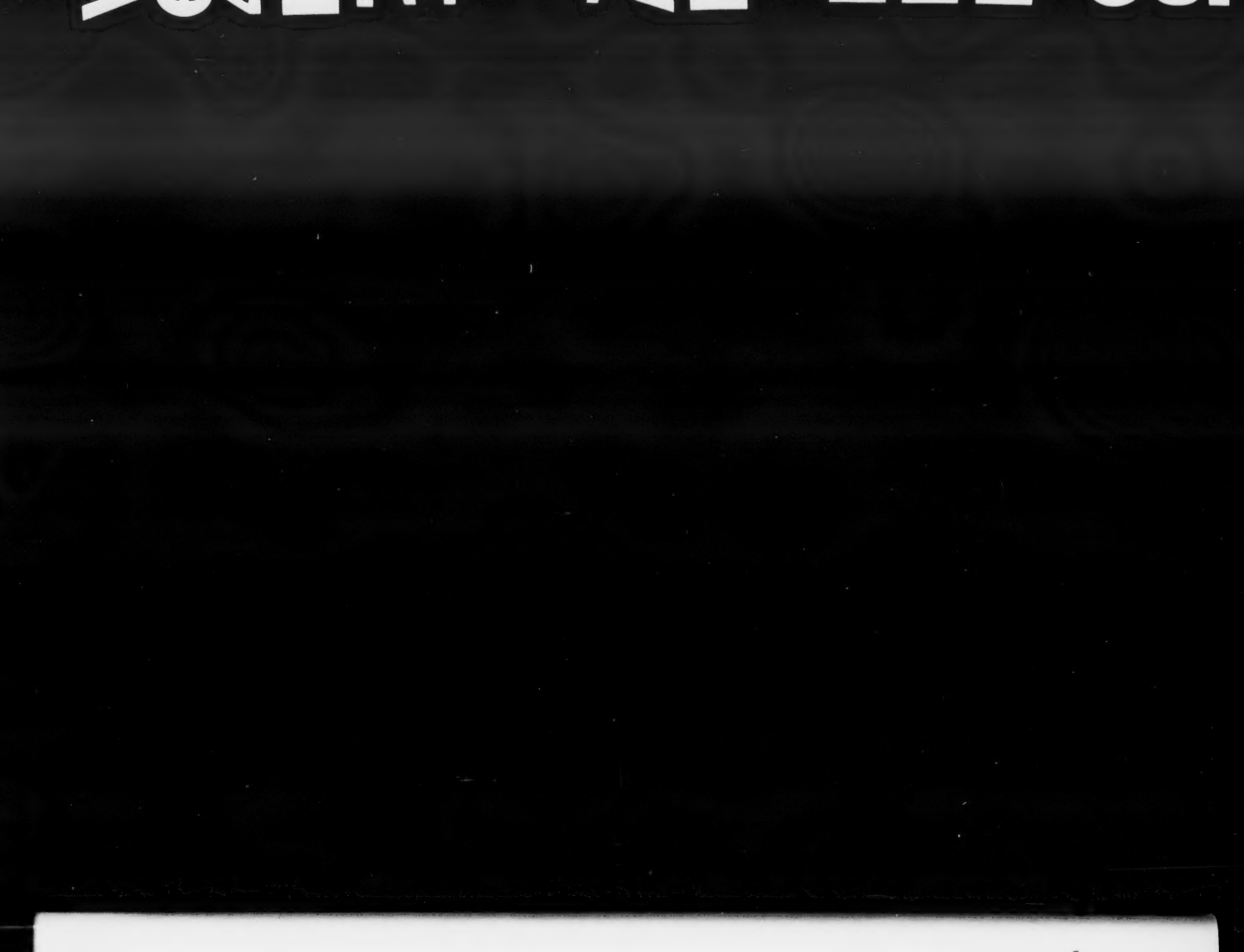
PENALTY FOR PRIVATE USE TO AVOID
PAYMENT OF POSTAGE, $300
(GPO)

National Bureau of Standards Journal of Research Vol. 64D, No. 1 January–February 1960
D. Radio Propagation

Contents

	Page
Effect of antenna size on gain, bandwidth, and efficiency. Roger F. Harrington.	1
Surface-wave resonance effect in a reactive cylindrical structure excited by an axial line source. A. L. Cullen.	13
Basic experimental studies of the magnetic field from electromagnetic sources immersed in a semi-infinite conducting medium. Martin B. Kraichman.	21
A very-low-frequency antenna for investigating the ionosphere with horizontally polarized radio waves. R. S. Macmillan, W. V. T. Rusch, and R. M. Golden.	27
Effects of high-altitude nuclear explosions on radio noise. C. A. Samson.	37
Measured frequency spectra of very-low-frequency atmospherics. Tatsuzo Obayashi.	41
Determination of the amplitude-probability distribution of atmospheric radio noise from statistical moments. W. Q. Crichlow, C. J. Roubique, A. D. Spaulding, and W. M. Beery.	49
Measurements of coastal deviation of high-frequency radio waves. C. W. McLeish.	57
An exact earth-flattening procedure in propagation around a sphere. Benjamin Y.-C. Koo and Martin Katzin.	61
Limit of spatial resolution of refractometer cavities. William J. Hartman.	65
Conference on arctic communication. R. C. Kirby and C. G. Little.	73
Tropospheric scatter propagation and atmospheric circulations. William F. Moler and D. B. Holden.	81
Layered earth propagation in the vicinity of Point Barrow, Alaska. Glenn M. Stanley.	95
Announcement of systematic ionospheric electron density data.	99
Publications of the National Bureau of Standards.	101

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. Price 70 cents (single copy). Subscription price: $4.00 a year; $0.75 additional for foreign mailing.
UNITED STATES GOVERNMENT PRINTING OFFICE, WASHINGTON : 1960



Selected Abstracts

Earth currents near a monopole antenna with symmetrical top loading, James R. Wait, *J. Research NBS* **62,** *247* (*June 1959*).

Expressions for the fields are developed for a vertical ground-based monopole with a cone or disk located at the top of the antenna to simulate umbrella top loading. The current distribution on the structure is assumed. Using spherical-wave functions, the magnetic-field distribution on the ground plane near the base of the antenna is computed and illustrated by graphs. For the case where the antenna is electrically small, the currents flowing on the cone or disk are shown to contribute only slightly to the total field.

Earth currents near a top-loaded monopole antenna with special regard to electrically small L- and T-antennas, H. Lottrup Knudsen, *J. Research NBS* **62,** *283* (*June 1959*).

An investigation has been made of the ground currents near a top-loaded monopole with nonazimuthal symmetry. Formulas have been developed for the surface current density produced by an inclined, straight wire over a horizontal ground plane for an arbitrary current distribution on the antenna. Working formulas have been developed and numerical calculations of the surface current density on the ground plane have been carried out for the case of a small antenna with a linear current distribution. These results have been used for the calculation of the contribution to the surface current density due to the top loading in the case of an L-antenna and in the case of a T-antenna. In each case both the absolute value of the surface current density arising from the top loading and the relative value of its ϕ-component have been plotted. This component under certain circumstances may be important in calculating the ground losses in the case of a system of radial ground wires.

World maps of F2 critical frequencies and maximum usable frequency factors, D. H. Zacharisen, *NBS 2 Tech. Note* (*PB151361*) *$3.50*.

This report was prepared for the purpose of presenting six months of contour maps and charts for use in predicting F2-layer maximum usable frequencies. Prediction maps for each even hour of Greenwich Mean Time and charts in which time is continuous along the abscissa are given for the months of January, March, June, July, September, and December. The four parameters used for predicting MUFs are foF2 and the 4000 km MUF factor for a twelve-month running average Zurich sunspot number of 50, and the rates of change of foF2 and 4000 km MUF factor with sunspot number. The first three parameters use a map presentation with GMT constant over the surface of the map. The fourth parameter uses a chart presentation in which the ordinate is geomagnetic latitude and the abscissa is local time.

Low- and very-low-radiofrequency tables of ground-wave parameters for the spherical earth theory: the roots of Riccati's differential equation (*supplementary numerical data for NBS Circ. 573*), J. R. Johler, L. C. Walters, C. M. Lilley, *NBS Tech. Note* (*PB151366*) *$2.25*.

The roots of Riccati's differential equation are tabulated in detail for the low- and very-low radiofrequencies. The zeroes and certain other parameters used in the calculation of the amplitude and phase of the ground wave by the conventional "residue series" are also tabulated.

Green coronal line intensity and geomagnetism, C. Warwick, *J. Geophys. Research* **64,** *527* (*May 1959*).

Climax Observatory spectrograms of the green coronal line, λ5303, for the period 1942 to 1944 near solar activity minimum have recently been remeasured. Analysis of these revised intensities indicates that the most consistent feature in the relation of green line intensity to geomagnetism is a minimum in geomagnetic activity following central meridian passage (CMP) of regions of high green line intensity. This effect may be interpreted as a result of the tendency of solar activity regions to coincide with regions of high green line intensity and to avoid M regions. In the period October 1953 to October 1954, at sunspot minimum, no relation appeared between green line intensity and geomagnetic activity.

Calculated patterns of slotted elliptic-cylinder antennae, James R. Wait and Walter E. Mientka, *Appl. sci. Research, Sec. B* **7,** *449* (*1959*).

The model assumed is a perfectly conducting elliptic cylinder of infinite length which has a narrow axial slot of finite length. Patterns are presented for various ratios of the major to minor axis. It is indicated that for reasonably large elliptic cylinders the patterns are dependent mainly on the surface curvature in the neighborhood of the slot.

A study of limb flares and associated events, C. Warwick and M. Wood, *Astrophys. J.* **129,** *801* (*May 1959*).

Measured heights of limb flares were used to derive a frequency distribution of flare height and to assess the importance of height in the production of flare-associated events. The observed height distributions can be explained by a real distribution with maximum at zero, decreasing toward greater heights. The occurrence of SWF (short-wave fadeout) depends strongly on flare height. This dependence could be explained as an effect either of absorption of the ionizing radiation in the sun's atmosphere or of a mechanism of production of ionizing radiation that depends critically on the height of the flare. Center-to-limb variation of SWF occurrence, while not conclusive, favors the former interpretation. The relation between optical flare characteristics and occurrence of associated bursts of radionoise is strongest for decimeter bursts and weaker for bursts at meter wave lengths.

Some characteristics of VLF propagation using atmospheric waveforms, William L. Taylor and L. Jerome Lange, *Recent advances in atmospheric electricity, p. 609* (*Pergamon Press, 1959*).

Simultaneous observations have been made of the waveforms of atmospherics ("sferics") at four widely separated stations. VLF attenuation characteristics have been computed by comparing the spectra of the waveforms recorded at the four locations for each discharge. The results of some preliminary measurements of VLF attenuation are presented for the band of frequencies from 4 kc to 30 kc at distances of 1,200 km to 6,500 km.

Other NBS Publications*

Journal of Research, Section 63A, No. 3, November–December 1959. 70 cents.

Multiple ionization of rare gases by electron impact. M. Krauss, R. M. Reese, and V. H. Dibeler.

Light scattering by commercial sugar solutions. Carl J. Rieger and Frank G. Carpenter.

Analysis of the first spectrum of ruthenium (Ru I). K. G. Kessler.

Supplementary Zeeman data for the first spectrum of ruthenium (Ru I). J. Rand McNally, Jr., and Karl G. Kessler.

Low even configurations in the first spectrum of ruthenium (Ru I), part 2. R. E. Trees.

Thermal degradation of polymers at high temperatures. Samuel L. Madorsky and Sidney Straus.

Influence of impurities on the pyrolysis of polyamides. Sidney Straus and Leo A. Wall.

A preliminary list of levels and *g*-values for the first spectrum of thorium (Th I). Romuald Zalubas.

OH in the solar spectrum. Charlotte E. Moore and Herbert P. Broida.

Use of Chebychev polynomials in thin film computations. Klaus D. Mielenz.

World maps of F2 critical frequencies and maximum usable frequency factors, D. H. Zacharisen, NBS Tech. Note 2 (PB151361) $3.50 (see abstracts above).

Low- and very low-radiofrequency tables of ground wave parameters for the spherical earth theory: The roots of Riccati's differential equation. (Supplementary numerical data for NBS Circ. 573) J. R. Johler, L. C. Walters, and C. M. Lilley, NBS Tech. Note 7 (PB151366) $2.25 (see abstracts above).

Calculated behavior of a fast neutron spectrometer based on the total absorption principle, J. E. Leiss, NBS Tech. Note 10 (PB151369). $1.00.

Penetration of gamma rays from isotropic sources through aluminum and concrete, M. J. Berger and L. V. Spencer, NBS Tech. Note 11 (PB151370). 50 cents.

Prediction of the cumulative distribution with time of ground wave and tropospheric wave transmission loss. Part 1. The prediction formula, P. L. Rice, A. G. Longley, and K. A. Norton, NBS Tech. Note 15 (PB151374). $1.50.

The changing character of chemical research in government, E. Wichers, The Chemist XXXVI, No. 7, 260 (1959).

Study of the setting of plaster, K. D. Jorgensen and A. S. Posner, J. Dental Research **38,** No. 3, 491 (1959).

Radiation attenuation data, H. O. Wyckoff, Radiation Hygiene Handbook, edited by H. Blatz, Section 8, p. 1 (McGraw-Hill Book Co., New York, N.Y., 1959).

Corrosion of type 310 stainless steel by synthetic fuel oil ash, H. L. Logan, Corrosion **15,** 443t (1959).

Physical research, Part 2, G. C. Paffenbarger and W. Souder, J. Am. Dental Assoc. **58,** 98 (1959).

Effects of gamma radiation on collagen, J. Cassel, J. Am. Leather Chemists' Assoc. LIV, No. 8, 432 (1959).

Standards for neutron flux measurement and neutron dosimetry, R. S. Caswell, E. R. Mosburg, Jr., and J. Chin, 2d United Nations International Conf. on the Peaceful Uses of Atomic Energy, Vol. 21, Health and Safety: Dosimetry and Standards, P/752 USA, p. 92 (1959).

Note on bivariate linear interpolation for analytic functions, W. Gautschi, Math. Tables and Other Aids to Computation **13,** No. 66, 91 (1959).

Absorption of radiation by a cylindrical sample of a strong absorber, P. H. Fang and I. A. Stegun, Letter to Editor, J. Chem. Phys. **31,** 267 (1959).

PILOT—A new multiple computer system, A. L. Leiner, W. A. Notz, J. L. Smith, and A. Weinberger, J. Assoc. Computing Mach. **6,** No. 3, 313 (1959).

Properties of rutile (titanium dioxide), F. A. Grant, Rev. Modern Phys. **31,** 646 (1959).

Analysis of vibrational relaxation data in shock wave experiments, K. E. Shuler, J. Chem. Phys. **30,** No. 6, 1631 (1959).

The interlaboratory evaluation of testing methods, J. Mandel and T. W. Lashof, ASTM Bull. No. 239, 53 (1959).

Branched-chain higher sugars. I. A 9-aldo-4-C-formyl-nonose derivative, R. Schaffer and H. S. Isbell, J. Am. Chem. Soc. **81,** 2178 (1959).

The structure of electrolytic solutions, W. J. Hamer (John Wiley & Sons, Inc., New York, N.Y., 1959).

The use of a vacuum microbalance in studies of electron tube materials, G. F. Rouse, Proc. 4th Natl. Conf. on Tube Techniques, September 10–12, 1958, Sponsored by the Advisory Group on Electron Tubes, p. 262 (New York University Press, New York, N.Y., 1959).

Widths of cracks in concrete at the surface of reinforcing steel evaluated by means of tensile bond specimens, D. Watstein and R. G. Mathey, J. Am. Concrete Inst. **31,** No. 1, 47 (1959).

Exposure standards and radiation protection regulations, L. S. Taylor, Radiation Hygiene Handbook, edited by H. Blatz, Section 3, p. 2 (McGraw-Hill Book Co., New York, N.Y., 1959).

Geometrical anisotropy of magnetic materials in wave guides and cavities, A. L. Steinert, J. Appl. Phys. **30,** 1109 (1959).

Radioactivity standardization in the United States, W. B. Mann and H. H. Seliger, 2d U.N. Intern. Conf. on the Peaceful Uses of Atomic Energy, Vol. 21, Health and Safety: Dosimetry and Standards, P/750 USA, p. 90 (1959).

Magnetic interaction of H_3, V. Griffing, J. L. Jackson, and B. J. Ransil, J. Chem. Phys. **30,** No. 4, 1066 (1959).

Program of the International Commission on radiological units and measurements, L. S. Taylor, L. H. Gray, and H. O. Wyckoff, 2d U.N. Intern.

Conf. on the Peaceful Uses of Atomic Energy, Vol. 21, Health and Safety: Dosimetry and Standards, P/2243 WHO, p. 81 (1959).

New approach in the theory of satellite orbits, J. P. Vinti, Phys. Rev. Letters **3,** No. 1, S517 (1959).

The nature, cause and effect of the porosity in electro-deposits, IV. Influence of gas bubbles on the formation of pores, F. Ogburn and D. W. Ernst, Plating **46,** 957 (1959).

Low even configurations in the first spectrum of ruthenium (Ru I), R. E. Trees, J. Opt. Soc. Am. **49,** 838 (1959).

End plate modification of X-band TE_{011} cavity resonators, M. C. Thompson, Jr., F. E. Freethey, and D. M. Waters, Letter to editor, IRE Trans. on Microwave Theory Tech. **MTT–7,** 388 (July 1959).

**Publications for which a price is indicated (except for Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (order by PB number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March-April 1960

Optimum Frequencies for Outer Space Communication[1]

George W. Haydon[2]

(November 10, 1959)

Frequency dependence of radio propagation and other technical factors which influence outer space communications are examined to provide a basis for the selection of frequencies for communication between earth and a space vehicle or for communication between space vehicles.

1. Introduction

The probable future rapid advances in the use of satellites and space vehicles will intensify the requirements for adequate space communications. Since only modest transmitter power will be initially available in the space vehicle, careful engineering of the space circuits will be necessary to assure adequate communications and particular attention to the selection of radiofrequencies will be required. Optimum frequencies can be selected on the basis of the signal-to-noise ratio for a given transmitter power, a minimum distortion of phase and amplitude, the minimum likelihood of interference from other equipment, etc. This report takes signal-to-noise ratio as the sole criteria of frequency selection recognizing that diffraction and other distortions may cause problems in tracking and location.

2. Factors Affecting the Selection of Frequencies

All communication between earth and outer space must pass through the earth's atmosphere (including the troposphere and ionosphere). Communication between satellites will primarily involve radio paths outside the influence of the earth's atmosphere.

The atmosphere is frequency selective, allowing some frequencies to pass through readily while severely attenuating others. A range of frequencies in which waves readily penetrate the atmosphere is often called a "window."

Two principal ranges of frequencies pass readily through the atmosphere. They are: (1) The range between ionospheric critical frequencies and frequencies absorbed by rainfall and gases (about 10 to 10,000 Mc), and (2) the combined visual and infrared ranges (about 10^8 to 10^9 Mc).

The atmosphere is known to be partially transparent in a third range below about 300 kc. Waves are propagated through the ionosphere in this range by what is sometimes called the whistler mode. Propagation in this mode is not yet well understood.

The range 10 to 10,000 Mc is the most practical for communication purposes considering the present state of development in radiofrequency power generation. The upper limit of this range may be as low as 5,000 to 6,000 Mc during heavy rainstorms and the lower limit may be as high as 80 to 100 Mc depending upon the degree of solar activity, the location of the earth terminal, and the geometry of the signal path. On the other hand, the window may extend from as low as 2 Mc for polar locations during night-time periods to as high as 50,000 Mc at high altitude rain-free locations.

In the midportion of this window favorable propagation conditions exist, and circuit performance can be estimated on the basis of free-space propagation conditions by the following formula:

$$P_t' \propto \left(\frac{P_r' f^2 d^2}{G_t G_r}\right)$$

where: P_t'=required transmitter power, P_r'=minimum permissible receiver input power, f=frequency, d=distance between transmitter and receiver, G_t=transmitting antenna gain power, G_r=receiving antenna gain power.

Actual propagation conditions vary substantially from this free space assumption at frequencies near the edge of the radio window, and it is necessary to correct for ionospheric and tropospheric effects to obtain a true estimate of frequency dependence. This correction requires an estimate of tropospheric absorption [1][3] at the higher frequencies and an estimate of ionospheric absorption at the lower frequencies [2]. In addition to estimating ionospheric absorption, an estimate of the probability of radio signals penetrating the ionosphere must be made [3].

To determine optimum frequencies, the variation of background radio noise within the radio window must also be considered:

(1) Cosmic noise predominates at the lower edge of the radio window and decreases with frequency until noise within the receiving equipment predominates.

(2) In most present-day facilities the receiving equipment noise tends to predominate above about 100 to 200 Mc for antennas directed toward average sky noise areas and above about 300 to 500 Mc for antennas directed toward high cosmic noise areas such as the Milky Way.

(3) If low noise receiving equipment such as the MASER amplifier is used, receiver noise may predominate above about 600 to 1,000 Mc.

[1] The basic material in this paper was unanimously adopted by the International Radio Consultative Committee at the IX Plenary Assembly in Los Angeles April 1959, and is being issued as CCIR Report No. 115, Factors affecting the selection of frequencies for telecommunication with and between space vehicles.

[2] United States Army Radio Frequency Engineering Office, Office of Chief Signal Officer, The Pentagon, Washington, D.C.; now with Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[3] Figures in brackets indicate the literature references at the end of this paper.

(4) Noise within conventional receivers normally increases slowly as the operating frequency is increased, but may tend to decrease at the higher frequencies if MASER amplifiers are employed.

For antennas of fixed physical size, high frequencies have the advantage of greater gain but the disadvantage of narrow beamwidths and associated tracking problems.

High speed vehicles traveling so that the distance between transmitter and receiver is rapidly changing have apparent frequencies differing from the actual transmitter frequencies by the Doppler frequency shift component in the direction of reception.

Within the solar system there is evidence of appreciable densities of electrons out to great distances from the sun.

Transmission time delay will become substantial in outer space communications, e.g., 2.6 sec are required for a round trip radio signal to the moon. This time delay is essentially independent of operating frequency.

3. Discussion

Although great distances are involved, the propagation medium in space beyond the first 500 miles of the earth's atmosphere is believed to be essentially transparent to radio waves. Thus we may estimate performance on the basis of free-space propagation. Frequency dependence of receiver input power under free-space propagation conditions depends upon the type of antenna at the transmitter and receiver. This frequency dependence is shown by the following free-space propagation formula:

$$P_r \propto \left(\frac{P_t G_t G_r}{f^2 d^2}\right)$$

where: P_r is receiver input power, P_t is transmitter power, and other symbols have the same meaning as before.

Frequency dependence of receiver input power for free space propagation conditions can be summarized as follows:

(1) If both the transmitting and receiving terminals of a free-space communication link use nondirective antenna (e.g., two vehicles in space) or if beamwidths at both terminals are fixed, the receiver input signal power increases as the frequency is decreased:

$$P_r \propto \left(\frac{P_t}{f^2 d^2}\right).$$

(2) If one terminal of a free-space communication link uses a directive antenna of fixed physical size and can operate with narrower and narrower beamwidths as frequency increases and the other terminal uses a nondirective antenna or a fixed beamwidth antenna, e.g., a directive antenna on the earth's surface ($G_t \propto f^2$) and a nondirective antenna on a space vehicle, the receiver input signal power is independent of frequency [$P_r \propto (P_t/d^2)$].

(3) If both terminals of a free-space communication link use directive antennas of fixed physical size and can operate with narrower and narrower beamwidths as frequency increases, e.g., a directive antenna on the earth's surface and a directive antenna on a more elaborate space vehicle ($G_t \propto f^2$ and $G_r \propto f^2$), the receiver input signal power increases as the frequency is increased [$P_r \propto (P_t f^2/d^2)$].

Frequency dependence for practical space-communication circuits requires that atmospheric effects be included. Receiver input signal power and receiver input noise power for a directive receiving antenna and nondirective transmitting antenna are shown in figure 1. The receiver input power includes ionospheric and tropospheric effects for a 1,000-mile propagation path tangential to the earth's surface for summer midday operation during periods of high solar activity and for moderate rain conditions such as experienced 1 percent of the time in the Washington, D.C., area. This is typical of the most adverse propagation conditions normally encountered in the absence of sudden ionospheric disturbances, instances of intense sporadic E, areas of auroral activity, or rain conditions of cloudburst proportions. During more favorable propagation conditions, such as a propagation path normal to the earth's surface during the night at the lower frequencies or in a high altitude rain-free location for the higher frequencies, the receiver input power can be expected to be essentially independent of frequency over a wider range of frequencies. The receiver input power as shown between 100 and 500 Mc in figure 1 will be typical over a much wider frequency range during these favorable propagation periods.

Figure 2 shows essentially the same information as figure 1, except that the distance is increased to 300,000 miles, the receiving antenna diameter is increased to 120 ft, the use of cooled amplifiers such as the MASER is anticipated, quasi-maximum

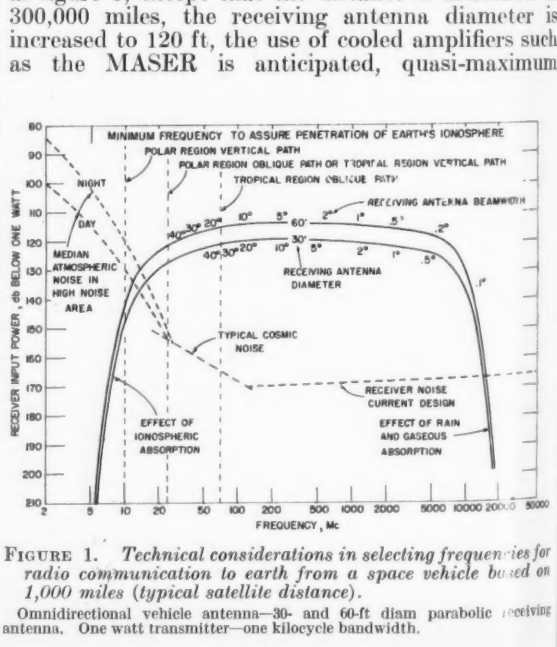

MINIMUM FREQUENCY TO ASSURE PENETRATION OF EARTH'S IONOSPHERE
POLAR REGION VERTICAL PATH
POLAR REGION OBLIQUE PATH OR TROPICAL REGION VERTICAL PATH
TROPICAL REGION OBLIQUE PATH
RECEIVING ANTENNA BEAMWIDTH
NIGHT
DAY
MEDIAN ATMOSPHERIC NOISE IN HIGH NOISE AREA
RECEIVING ANTENNA DIAMETER
TYPICAL COSMIC NOISE
RECEIVER NOISE CURRENT DESIGN
EFFECT OF IONOSPHERIC ABSORPTION
EFFECT OF RAIN AND GASEOUS ABSORPTION
RECEIVER INPUT POWER, db BELOW ONE WATT
FREQUENCY, Mc

FIGURE 1. *Technical considerations in selecting frequencies for radio communication to earth from a space vehicle based on 1,000 miles (typical satellite distance).*

Omnidirectional vehicle antenna—30- and 60-ft diam parabolic receiving antenna. One watt transmitter—one kilocycle bandwidth.

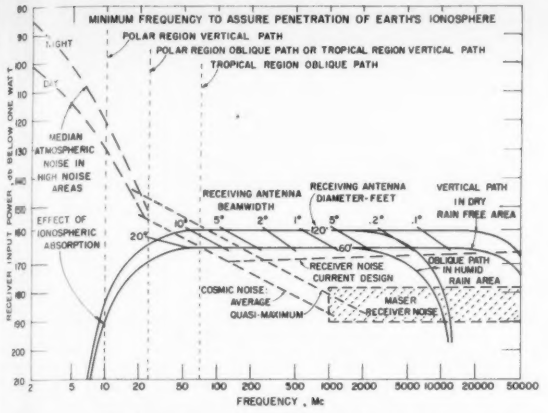

MINIMUM FREQUENCY TO ASSURE PENETRATION OF EARTH'S IONOSPHERE
POLAR REGION VERTICAL PATH
POLAR REGION OBLIQUE PATH OR TROPICAL REGION VERTICAL PATH
TROPICAL REGION OBLIQUE PATH
NIGHT
DAY
MEDIAN ATMOSPHERIC NOISE IN HIGH NOISE AREAS
EFFECT OF IONOSPHERIC ABSORPTION
RECEIVING ANTENNA BEAMWIDTH
RECEIVING ANTENNA DIAMETER-FEET
VERTICAL PATH IN DRY RAIN FREE AREA
OBLIQUE PATH IN HUMID RAIN AREA
RECEIVER NOISE CURRENT DESIGN
COSMIC NOISE AVERAGE QUASI-MAXIMUM
MASER RECEIVER NOISE
RECEIVER INPUT POWER, db BELOW ONE WATT
FREQUENCY, Mc

FIGURE 2. *Technical considerations in selecting frequencies for radio communication to earth from a space vehicle based on 300,000 miles (distance to circle moon).*

Omnidirectional vehicle antenna—60- and 120-ft diam parabolic receiving antenna. One watt transmitter—one kilocycle bandwidth.

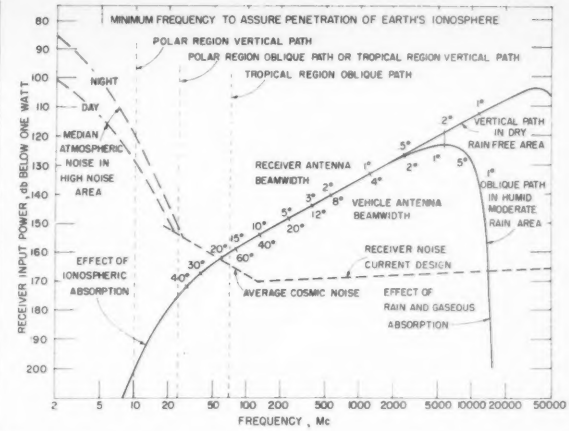

MINIMUM FREQUENCY TO ASSURE PENETRATION OF EARTH'S IONOSPHERE
POLAR REGION VERTICAL PATH
POLAR REGION OBLIQUE PATH OR TROPICAL REGION VERTICAL PATH
TROPICAL REGION OBLIQUE PATH
NIGHT
DAY
MEDIAN ATMOSPHERIC NOISE IN HIGH NOISE AREA
EFFECT OF IONOSPHERIC ABSORPTION
RECEIVER ANTENNA BEAMWIDTH
VEHICLE ANTENNA BEAMWIDTH
VERTICAL PATH IN DRY RAIN FREE AREA
OBLIQUE PATH IN HUMID MODERATE RAIN AREA
RECEIVER NOISE CURRENT DESIGN
AVERAGE COSMIC NOISE
EFFECT OF RAIN AND GASEOUS ABSORPTION
RECEIVER INPUT POWER, db BELOW ONE WATT
FREQUENCY, Mc

FIGURE 3. *Technical considerations in selecting frequencies for radio communication to earth from a space vehicle based on 300,000 miles (distance to circle moon).*

Fifteen foot diameter parabolic vehicle antenna—60-ft diam parabolic receiving antenna. One watt transmitter—one kilocycle bandwidth.

values of cosmic noise for high gain antennas are given, and receiver input power is shown for a vertical path in a dry rain-free location.

Figure 3 shows the theoretical improvement at the higher frequencies if a directive antenna is used on the space vehicle. Although receiver input power is shown only for a 15-ft diam parabolic antenna on the space vehicle, this improvement with increased frequency applies for all directive antennas of fixed physical size. Since the increase in antenna gains at the higher frequencies more than offsets the slightly increased power requirement due to increased receiver noise at these frequencies, the first impression is that the higher the frequency the better the expected circuit performance as long as the frequency is below the upper limit of the radio window. (About 6,000 Mc for oblique paths during moderate rainfall and up to 50,000 Mc for high altitude rain-free location.) The physical problem of antenna design and tracking, however, establishes minimum permissible antenna beamwidths and is believed to place practical limits on this upper frequency at much lower values.

Theoretical effective power requirements for greater distances are shown in figure 4 as a function of frequency. Power requirements shown are the minimum detectable radiated power (6 db S/N ratio) from a space vehicle to a 60 ft diam earth-based antenna under the most adverse propagation conditions normally encountered. Allowances for fading and antenna beamwidth limitations are not shown by the chart. Approximate distances from earth to the moon, the sun, certain planets, and to typical manmade satellites are shown.

For systems with minimum permissible beamwidths, any increase in antenna size reduces this practical upper frequency since antenna beamwidth decreases with increase in antenna size. The relationship between antenna size, operating frequency, and antenna beamwidth for parabolic antennas is shown in figure 5. If either the ground terminal or

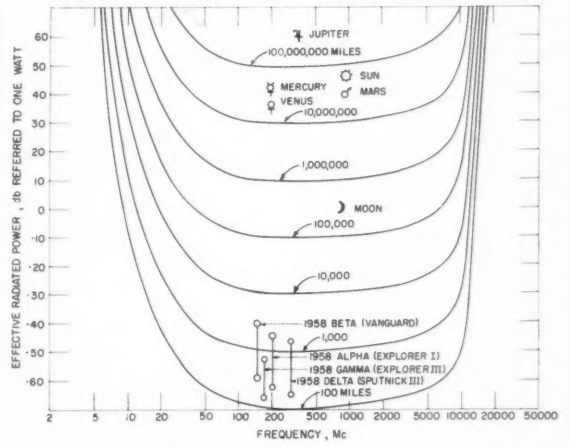


FIGURE 4. *Theoretical effective radiated power required from space vehicle to permit detection on earth.*

One kilocycle bandwidth—6 db signal-to-noise ratio—receiving antenna—60-ft parabola—daytime operation—high solar activity—rain and gaseous absorption based on 1 percent of time (Washington, D.C.). Radio path approximately horizontal to earth—receivers of current design.

space vehicle maximum antenna physical size and minimum antenna beamwidth is fixed, the optimum frequency can be estimated from figure 5.

Figure 6 is a nomogram to estimate the bandwidth allowance required to accommodate the Doppler frequency shift for radio transmissions from high speed vehicles.

For communication between vehicles in outer space, free-space propagation conditions apply over a wide frequency range. Frequencies above or below the earth's radio window can be expected to minimize interference problems with operations on earth. These frequencies will be below about 10 Mc or above about 10,000 Mc. Selection of an optimum frequency can be based on free-space

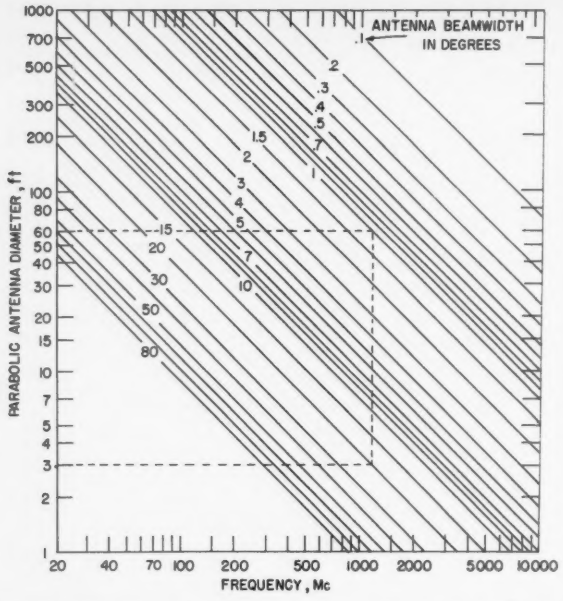


FIGURE 5. *Chart to estimate optimum frequencies for outer space communication when directive antennas are used at both terminals, and physical antenna size and minimum permissible antenna beamwidth are fixed factors.*

Example: Based on a 1 deg minimum beamwidth for the earth terminal, the optimum frequency for a 60-ft diam parabolic is about 1,200 Mc and the optimum vehicle antenna size is a 3-ft diam parabolic based on a 20-deg minimum beamwidth criteria for the space vehicle antenna.

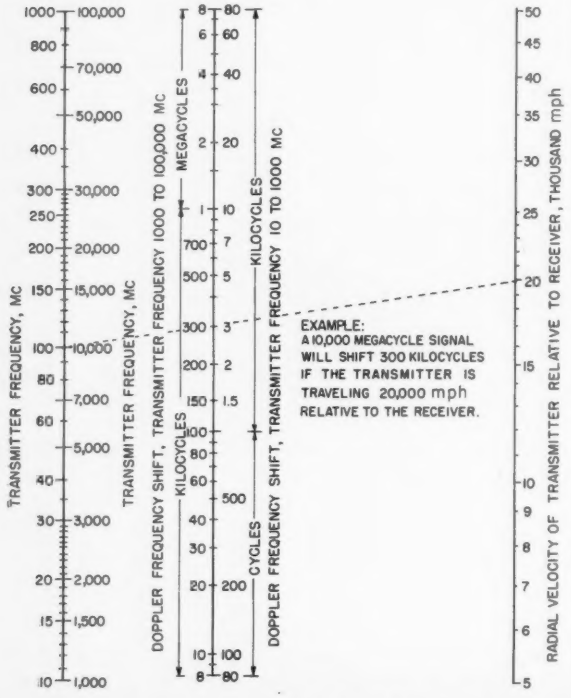


FIGURE 6. *Chart to estimate Doppler frequency shifts.*

propagation but requires an estimate of noise powers in outer space, particularly as to the radio noise at frequencies between 2 and 10 Mc. Frequencies below 2 Mc are considered impractical because of antenna sizes required and the substantial plasma frequencies probably occurring in outer space during periods of severe magnetic storms. If radio noise is excessive below 10 Mc and if antenna orientation problems limit the use of high gain antennas, the optimum frequency for communication between space vehicles may fall within the 10 to 10,000 Mc radio window.

4. Conclusions

Communications between earth and outer space are theoretically possible within two broad frequency bands; about 10 to 10,000 Mc and in the infrared and optical regions. At the current state of the equipment art the lower band is definitely favored.

The upper limit of the lower band is dependent upon tropospheric conditions, and the lower limit depends upon ionospheric conditions. The band, therefore, is not sharply defined but is dependent upon geographic location and time of operation. For reliable communication to any earth terminal location, this lower band narrows to about 70 to 6,000 Mc.

Within the 70- to 6,000-Mc band the optimum frequency will depend upon the specific communication service required and will be a compromise between the maximum practical antenna size, the minimum beamwidth which will permit acquisition or tracking, and the radio noise levels.

For space vehicles using essentially omnidirectional antennas communicating with earth terminals using directive antennas, the receiver input power will be constant with frequency over much of the 70- to 6,000-Mc band, and the background noise level and beamwidth requirements to assure tracking determines the optimum frequency. Background noise from sources within the antenna beam (cosmic noise) predominates at the lower edge of the band and noise generated within the first stages of the receiver predominates at the upper edge of the band. The crossover point between these noise sources determines the frequency with the maximum signal-to-noise ratio and, therefore, the optimum frequency for communication if antenna beamwidths are satisfactory at these frequencies. These optimum frequencies are about as follows:

(1) 100 to 200 Mc for conventional receivers with antennas directed toward average cosmic noise sources;

(2) 300 to 500 Mc for conventional receivers with high gain antennas directed toward high cosmic noise sources such as the Milky Way;

(3) 1,000 to 3,000 Mc if the receiver is equipped with cooled amplifiers such as the MASER.

Antenna beamwidths must always be considered and compromises made between beamwidth and optimum signal-to-noise ratios. Since receiver noise increases only slowly with frequency and

receiver input power is constant up to about 6,000 Mc, higher frequencies may be used with only slight decrease in S/N ratio but at the expense of increased tracking difficulty.

As more elaborate space vehicles capable of maintaining attitude and employing directive antennas are developed, the receiver input power will increase with frequency and the background noise level will no longer determine the optimum frequency. The optimum frequency will be governed by a compromise between maximum practical physical antenna size and the minimum antenna beamwidth consistent with acquisition and tracking techniques. If attitude control of the space vehicle and acquisition and tracking limitations of the ground stations establish minimum antenna beamwidths at both terminals, the fixing of the maximum practical antenna size at either terminal will establish the optimum frequency and antenna size for the other terminal. As attitude control and tracking techniques improve the optimum frequency increases. As larger antennas become practical the optimum frequency decreases. The optimum frequency is therefore closely associated with particular applications and can be selected once the physical antenna size and minimum beamwidths are established. For a 1-deg minimum beamwidth for the earth antenna and a 20-deg minimum beamwidth for the space vehicle antenna, optimum combinations of frequencies and antenna sizes are shown in table 1.

TABLE 1

Earth antenna diam	Optimum frequency	Space vehicle optimum antenna diam
ft	*Mc*	*ft*
30	2400	1½
60	1200	3
120	600	6

The optimum frequency for communication between outer space vehicles is unknown. It will depend upon radio noise in outer space and upon the ability of the vehicles to maintain attitudes and thereby use directional antennas. For omnidirectional antennas or for any fixed antenna beamwidth the optimum frequency will be the lowest frequency consistent with the practical antenna size and outer space radio noise levels. Since operation at frequencies outside the radio window will tend to minimize the radio interference problem between the space vehicles and earth, space vehicles with omnidirectional or broad beamwidth antennas should be assigned trial frequencies below 10 Mc if antennas at these frequencies are practical. For more elaborate space vehicles with the ability to properly orient antennas with very narrow beamwidths, operation at frequencies above or near the upper edge of the radio window is recommended (above 10,000 Mc). If physical antenna size limits the use of frequencies below 10 Mc and if the inability to orient antennas limits the use of frequencies above 10,000 Mc, antenna size and antenna beamwidth compromises will determine optimum frequencies. Frequencies may then be selected by the use of figure 5 in the same manner as for communication between space vehicles and earth, when both terminals use directive antennas except that the earth's ionospheric and tropospheric limitations will no longer apply.

5. References

[1] B. R. Bean and R. Abbott, Oxygen and water vapor absorption of radio waves in the atmosphere. Geofisica Pura E Appl. Milano **37,** pp. 127–144 (1957).

[2] P. O. Laitinen and G. W. Haydon, Analysis and predictions of sky wave fields intensity, Radio Propagation Agency Tech. Rept. No. 9. U.S. Army Signal Radio Propagation Agency, Fort Monmouth, N.J. (Aug. 1950).

[3] D. H. Zacharisen, World maps of *F*2 critical frequencies and maximum usable frequency factors, NBS Tech. Note No. 2 (PB151361) Apr. 1959.

The following paper has just been brought to the author's attention:

S. Perlman, L. C. Kelley, W. T. Russell, Jr., and W. D. Stuart, Concerning optimum frequencies for space communication, IRE Trans. on Communs. Systems, **CS–7,** 167 (Sept. 1959).

BOULDER, COLO.

(Paper 64D2–44)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March–April 1960

The Joint Use of the Ordinary and Extraordinary Virtual Height Curves in Determining Ionospheric Layer Profiles

L. R. O. Storey[1]

(October 8, 1959)

An extension of Budden's matrix method for determining ionospheric layer profiles is described. When analyzing vertical incidence ionograms by the matrix method, it is usual to interpret the virtual height curve for the ordinary mode only. Errors then arise from the presence in the lower ionosphere of low-density ionization for which the plasma frequency is less than the lowest frequency observed. In the proposed extension of the method, such errors are reduced by making use of the extraordinary virtual height curve as well as of the ordinary.

1. Introduction

In 1954 Budden [1][2] introduced the matrix method for analyzing the records of vertical incidence ionospheric soundings, and since then it has come into wide use. The records, which are known as "ionograms," contain curves of the virtual reflection heights of the ordinary and extraordinary modes as functions of frequency ("virtual height curves"), and the object of the analysis is to determine, from one or the other of these curves, how the ionization in the underside of the ionosphere is distributed in height. The difficulty lies in the integral equation that relates this distribution to the virtual height curve; in general, it is too complex to be solved analytically. The principle of Budden's method is to approximate the integral by a sum of discrete values of the integrand, thus converting the integral equation into a matrix equation, which may be solved by inverting the matrix.

In its customary form, the method makes use only of the ordinary virtual height curve. The ordinates of this curve are measured from the ionogram at selected frequencies, and then this set of ordinates is multiplied by a matrix to produce a set of the corresponding ordinates of the layer profile. The matrix multiplication is a simple repetitive process and so lends itself readily to performance by a digital computer. As a result, the matrix method is now widely favored for the routine analysis of ionospheric data.

Unfortunately, however, two rather restrictive conditions have to be met if the method is to work properly: First, the original virtual height curve should extend down to zero frequency; second, the ionosphere should consist of a single layer only, so that a given density of ionization occurs at just one height on its underside. Only under these conditions does the integral equation define a unique distribution.

Usually neither condition is met in practice. Most ionospheric sounders do not operate below about 1 Mc, while the ionosphere often appears to be stratified into several distinct layers. Then there are many possible distributions that could have given rise to the observed portion of the virtual height curve. In fact, these uncertainties are the main sources of error in the analysis of ionograms by the matrix method.

This paper deals only with the problem of how to reduce the errors that arise from ignorance of the virtual heights at low frequencies. Here the ionosphere still is assumed to consist of a single layer. The analysis appears capable of being extended to cover the case of multiple layers with intervening "valleys," but this development is not considered.

When the virtual height curve is missing at the low-frequency end, the usual practice is to assume that the unknown virtual heights are constant and equal to the value at the lowest frequency observed; this assumption implies that there is no ionization below the level of reflection for that frequency. If such "low-density ionization" is present in fact, then the group delay that it produces in the observed frequency range is attributed wrongly to time spent by the waves in traveling up to the ionosphere, so that the true heights are overestimated.

So long as attention is confined to a single virtual height curve, there is no way of detecting the low-density ionization or of isolating its effects in the observed frequency range. But these effects are different for the two modes of propagation; the group delays differ in amount and vary with frequency in different ways. This fact suggests that the low-density ionization could be studied by measuring both the ordinary and extraordinary virtual height curves and combining the data that they contain. In the present paper, extensions of the Budden matrix method are developed to do this.

The contents of the paper are arranged as follows: In section 2 terms are defined, and the present method for the interpretation of a single virtual height curve

[1] Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.; present address, Centre National de la Recherche Scientifique, Service d'Aeronomie, Observatoire de Meudon (S.-&-O.), Paris, France.

[2] Figures in brackets indicate the literature references at the end of this paper.

is reviewed briefly; the work reviewed is chiefly that of Budden [1], Rydbeck [2], and Schmerling [3]. Section 3, which contains the substance of the paper, shows how the matrix method may be extended to make joint use of both virtual height curves; the errors that may arise in the inferred distribution of ionization, due to random errors in measuring the virtual heights, are examined also. The analysis is illustrated in section 4 by a worked example, in which an artificial ionogram, computed for a known distribution of ionization, is interpreted by the new procedures. Finally, the whole analysis is discussed and criticized in section 5, and the conclusions of the paper are summarized in section 6.

Certain previous authors also have considered using the two virtual height curves in conjunction. The suggestion was made originally by Jackson [4], but his approach, which is one of trial and error, is not suitable for extensive use. Thomas [5], however, in a recent survey paper, has outlined some work by Titheridge [6], who appears to have adapted Jackson's method into a more convenient form. Titheridge's work, which has not been published at the time of writing this paper, may anticipate some of the results presented here.

2. Interpretation of a Single Virtual Height Curve

2.1. The Integral Equations

In this section the integral equations are derived that relate the virtual height curves for the ordinary and extraordinary modes to the distribution of ionization.

At vertical incidence, the virtual height of reflection h' is given as a function of the wave frequency f by the integral

$$h'(f)=\int_0^{z_R}\mu'(f,N)\,dz, \tag{1}$$

where z is height in general, z_R is the true height of reflection for the given frequency, and μ' is the group refractive index, which depends both on the frequency and on the local electron density N.

In the limit of very low wave frequencies, the virtual and true heights of reflection for both modes all tend to a single limiting value z_0, the height of the base of the ionosphere. From ground level up to this height the waves travel at the speed of light, which is to say that their group refractive indices are unity.[3] Thus (1) can be written as

$$h'(f)=z_0+\int_{z_0}^{z_R}\mu'(f,N)\,dz, \tag{2}$$

where the integral now embraces only the range of height in which free electrons are present.

The general expression for the virtual height is completed by introducing the relation between the density of electrons and the height. However, instead of the density itself, it is more convenient to work with the "plasma frequency" f_N. which is proportional to its square root. The function $z(f_N)$, in which plasma frequency is viewed as the independent and height as the dependent variable, will be referred to here as the "layer profile." The profile is a multivalued function, even for a single layer, since a given value of f_N, if it occurs at all, must do so at least twice, once on the upper surface of the layer and once on its underside. The relationship between the true heights of reflection and the layer profile may be understood by considering the condition for reflection. The condition for an ionized medium to reflect radio waves is that its plasma frequency should exceed a certain critical value that depends on the frequency of the waves and also on the mode of propagation. This value will be written as $f_R(f)$. It follows that waves of a certain frequency and mode, propagated up into the ionosphere from below, are reflected at the first level that they encounter at which f_N is equal to the appropriate value of f_R. Obviously, this is the lowest such level. Hence, the true height of reflection is given by

$$z_R=\text{minimum value of } z(f_R). \tag{3}$$

To summarize, the true height of reflection is a particular value of z, determined by f_R, which in turn is a function of f; this function is different for the two modes of propagation.

Since all reflection takes place from the underside of the ionosphere, only the profile of this part needs to be considered. Further, since the ionosphere has been assumed to consist of a single layer, the profile of its underside is a monotonic function. In this case, f_N can be taken as the variable of integration. Also, the group refractive index μ' will be expressed as a function of f_N instead of N. With these changes, the integral equation becomes

$$h'(f)=z_0+\int_0^{f_R}\mu'(f,f_N)\left(\frac{dz}{df_N}\right)df_N. \tag{4}$$

Here the equation is still quite general, in that it applies equally to both modes. The particular equations for the ordinary and extraordinary modes are to be obtained by considering their respective reflection conditions, that is to say, the different relationships between f_R and f. These relationships are graphed in figure 1.

The ordinary mode behaves very simply; the critical plasma frequency for reflection is equal to the wave frequency. Thus,

$$f_R=f \qquad \text{(ordinary mode)} \tag{5}$$

[3] Actually μ' is slightly greater than unity even below the ionosphere due to the refraction of the air, but the departure from unity is negligible in the present context.

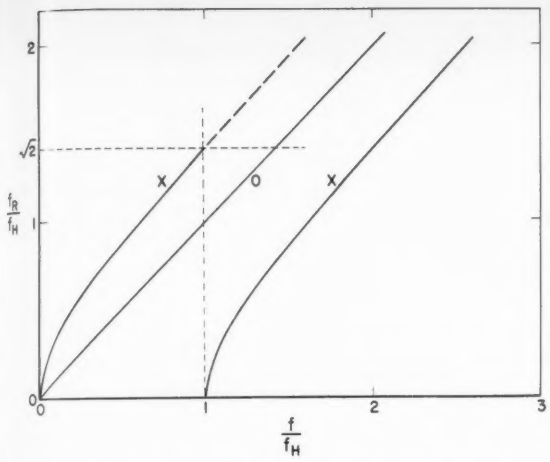


FIGURE 1. *The conditions for reflection of the ordinary* (O) *and extraordinary* (X) *magneto-ionic modes at vertical incidence.*

Abscissa: wave frequency (f). Ordinate: plasma frequency at the level of reflection (f_R). Both frequencies are normalized with respect to the gyrofrequency (f_H).

so the integral equation for the ordinary mode is

$$h'(f)=z_0+\int_0^f \mu'(f,f_N)\left(\frac{dz}{df_N}\right)df_N. \qquad (6)$$

From here on, the quantities h, h', and μ', when written without subscripts, will refer to the ordinary mode only. The corresponding quantities for the extraordinary mode will bear the subscript "x."

The extraordinary reflection condition is more complex. The relationship between f_R and f is now

$$f_R=\{f(f\pm f_H)\}^{\frac{1}{2}} \qquad \text{(extraordinary mode)} \quad (7)$$

$$=\tilde{f}, \qquad \text{say} \qquad (8)$$

where f_H is the gyrofrequency. The positive sign applies if $f<f_H$ (L.H. curve in fig. 1), while the negative sign applies in the more usual situation, where $f>f_H$ (R.H. curve).

Note that, as the wave frequency approaches the gyrofrequency from below, f_R tends to the value $\sqrt{2}$ f_H. On the other hand, as the gyrofrequency is approached from above, f_R becomes vanishingly small, so that the extraordinary reflection level should descend to the base of the ionosphere; actually, such echoes are never detected, because they are heavily absorbed.

The integral equation for the extraordinary mode can be written now as

$$h'_x(f)=z_0+\int_0^{\tilde{f}}\mu'_x(f,f_N)\left(\frac{dz}{df_N}\right)df_N, \qquad (9)$$

where the symbol $\tilde{f}$ has been adopted to denote the value of f_R for an extraordinary wave of frequency f.

This equation differs slightly from the corresponding equation for the ordinary mode, in that the upper limit of the integral ($\tilde{f}$) is not the same as the independent variable (f) on the left-hand side; also, to each value of $\tilde{f}$ below the limit $\sqrt{2}f_H$ there correspond two possible values of f (see fig. 1), one of which is less than the gyrofrequency, while the other is greater. However, only the upper range of f will be considered here, because most records of ionospheric virtual height do not extend below the gyrofrequency. With this restriction, there is a straightforward correspondence between $\tilde{f}$ and f, and the virtual height may be regarded equally well as a function of either variable. When it is regarded as a function of $\tilde{f}$, the extraordinary virtual height will be represented by the symbol $\tilde{h}$; thus

$$\tilde{h}(\tilde{f})=h'_x(f). \qquad (10)$$

To find the virtual height for a specified value of $\tilde{f}$, it is necessary to know the corresponding value of f. This frequency is given by the positive solution of the quadratic eq (7), and it is

$$f=\tfrac{1}{2}f_H\left\{\{1+4(\tilde{f}/f_H)^2\}^{\frac{1}{2}}\mp 1\right\}, \qquad (11)$$

where the positive sign now yields the range of frequencies above f_H, which is the range to be considered.

In these terms, the integral equation for the extraordinary mode is

$$\tilde{h}(\tilde{f})=z_0+\int_0^{\tilde{f}}\mu'_x(f,f_N)\left(\frac{dz}{df_N}\right)df_N. \qquad (12)$$

This equation is similar to (6), the only remaining difference being that, in calculating μ'_x, it is necessary first to obtain from (11) the value of f that corresponds to the given value of $\tilde{f}$.

The integral eqs (6) and (12) relate the virtual height curves to the slope of the layer profile. Were it possible to invert either equation so as to obtain a solution for the slope, then the profile itself could be recovered by a straightforward integration.

2.2. The Matrix Equations

The integral equations cannot be solved analytically in the general case, because the group refractive indices are complicated functions of f and f_N. Therefore, they are solved by converting them into matrix equations as will be described now.

Consider the eq (6) for the ordinary mode. The first step is to replace the continuous functions h' and z by their values at a number of discrete "sample frequencies." These frequencies will be written, in order of ascending magnitude, as f_0, f_1, f_2 . . . etc., or generally as f_m, where the index m is zero or a

positive integer; f_0 is zero frequency, and the rest are assumed to be spaced closely, but not necessarily evenly, through the full frequency range. The following notation will be used also:

$$\Delta f_m = f_m - f_{m-1} \qquad (m \geqq 1) \tag{13}$$

$$h'_n = h'(f_n) \tag{14}$$

$$z_m = z(f_m) \tag{15}$$

$$\Delta f_m = \begin{cases} z_m - z_{m-1} & (m \geqq 1) \\ z_0 & (m = 0). \end{cases} \tag{16}$$

The quantities Δz_m are increments of height on the layer profile; note the definition for $m=0$.

The next step in the approximation is to assume that the slope of the layer profile is effectively constant over each elementary frequency interval Δf_m, i.e.,

$$\frac{dz}{df_N} \simeq \frac{\Delta z_m}{\Delta f_m} \qquad (f_{m-1} \leqq f_N \leqq f_m).$$

Then the complete integral can be split up into a series of discrete sums, as follows:

$$h'_n \simeq \Delta z_0 + \sum_{m=1}^{n} \frac{\Delta z_m}{\Delta f_m} \int_{f_{m-1}}^{f_m} \mu'(f_n, f_N)\, df_N$$

$$\simeq \sum_{m=0}^{n} M_{nm} \Delta z_m, \qquad say \tag{17}$$

where Δz_0 has been written in place of z_0, and

$$M_{nm} = \begin{cases} 1 & (m = 0) \\ \dfrac{1}{\Delta f_m} \displaystyle\int_{f_{m-1}}^{f_m} \mu'(f_n, f_N)\, df_N & (1 \leqq m \leqq n) \\ 0 & m > n). \end{cases} \tag{18}$$

Now suppose that in (17), n is set equal in turn to 0, 1, 2, etc., up to the value that corresponds to the highest sample frequency. The result is an array of simultaneous equations, which may be solved for the unknowns Δz_m by the simple algebraic process that is described below in section 2.4. Originally, however, Budden [1] adopted an alternative approach, which is to regard the complete array as a single "matrix equation;" this approach has provided a more convenient basis for the subsequent extensions of Budden's methods. In matrix terminology, the set of quantities h'_n and also the set of Δz_m, both constitute "vectors," while the coefficients M_{nm} form a "lower triangular matrix." This differs from the corresponding matrix used by Budden (1955) and Schmerling (1957) by including an extra column with $m=0$, in which all the coefficients are unity, and an extra row with $n=0$, in which all the coefficients except the first are zero. Each of the remaining nonzero coefficients is the average of the values of group refractive index assumed by a wave of frequency f_n in traveling from the level where $f=f_{m-1}$ up to that where $f_N=f_m$.

If n is understood to run through its full range of values, then eq (17) states that the vector of the h'_n is obtained by multiplying the vector of the Δz_m by the matrix of the coefficients M_{nm}. Thus (17) is the required matrix equation for the ordinary mode.

The integral equation for the extraordinary mode can be treated likewise, and the result is the matrix equation

$$\tilde{h}_n \simeq \sum_{m=0}^{n} \tilde{M}_{nm} \Delta z_m \tag{19}$$

where

$$\tilde{M}_{nm} = \begin{cases} 1 & (m = 0) \\ \dfrac{1}{\Delta \tilde{f}_m} \displaystyle\int_{\tilde{f}_{m-1}}^{\tilde{f}_m} \mu'_x(f_n, f_N)\, df_N & (1 \leqq m \leqq n) \\ 0 & (m > n) \end{cases} \tag{20}$$

and the other new quantities are defined like their ordinary counterparts (cf. eqs (13), (14)).

From here on, the complete matrices will be represented by symbols in square brackets, while the same symbols without brackets will refer to the individual elements. Thus the group refractive index matrices will be written as $[M_{nm}]$ and $[\tilde{M}_{nm}]$. Their elements M_{nm} and $\tilde{M}_{nm}$ may be computed from the formulas given in appendix A.

2.3. Solution by Matrix Inversion

The matrix eqs (17) and (19) for the ordinary and extraordinary modes, respectively, may be solved by inverting the matrices. Thus, for the ordinary mode, the solution is

$$\Delta z_n \simeq \sum_{m=0}^{n} N_{nm} h'_m, \tag{21}$$

where the matrix $[N_{nm}]$ is the inverse of $[M_{nm}]$. Similarly, for the extraordinary mode,

$$\Delta z_n \simeq \sum_{m=0}^{n} \tilde{N}_{nm} \tilde{h}_m \tag{22}$$

where $[\tilde{N}_{nm}]$ is the inverse of $[\tilde{M}_{nm}]$. Because the matrices $[M_{nm}]$ and $[\tilde{M}_{nm}]$ are lower triangular, so also are their inverses.

The increments of height may be summed now to give the heights themselves:

$$z_n = \sum_{m=0}^{n} \Delta z_m. \tag{23}$$

This summation is analogous to the final integration that would be performed if the original integral equations, that involve the slope dz/df_N, could be solved analytically (see sec. 2.1).

2.4. Solution Row-by-Row

As an alternative to inverting the matrices, the increments of height on the layer profile may be found by solving the matrix equation row-by-row [7]. For instance, consider the eq (17) for the ordinary mode. It can be rearranged to read

$$\Delta z_n \simeq \frac{1}{M_{nn}}\left\{h'_n - \sum_{m=0}^{n-1} M_{nm}\Delta z_m\right\}. \qquad (24)$$

There is one such equation for each value of n. The complete set may be solved by starting from $n=0$ and working upwards. Thus,

$$\Delta z_0 \simeq \frac{1}{M_{00}}\{h'_0\},$$

$$\Delta z_1 \simeq \frac{1}{M_{11}}\{h'_1 - M_{10}\Delta z_0\},$$

$$\Delta z_2 \simeq \frac{1}{M_{22}}\{h'_2 - M_{21}\Delta z_1 - M_{20}\Delta z_0\},$$

and so on. In effect, the method consists of isolating, at each stage, the group delay produced in the unknown uppermost increment of height, by subtracting from the observed virtual height the group delay produced in the known height increments at lower levels.

The eq (19) for the extraordinary mode can be solved similarly. Matrix equations can be solved in this way only when, as in the present case, the matrices are triangular.

3. Joint Use of Both Virtual Height Curves

3.1. General Remarks

In the previous section, each of the matrix eqs (17) and (19) represented a set of linear simultaneous equations in which the number of data (ordinates of the virtual height curve) was equal to the number of unknowns (increments of height on the layer profile). Hence each set had an exact solution, so that the complete virtual height curve for either mode was sufficient by itself to determine the profile. But when some of the curve is missing, as when the records are unavailable at low frequencies, then the remaining high frequency data are insufficient to determine the profile without making special assumptions (see sec. 1).

However, the situation is not hopeless, since the data for the two magneto-ionic modes are mutually independent. Suppose, therefore, that in sampling the virtual height curves to obtain these data, the sample frequencies are chosen so that the unknowns are the same in the two sets of equations. Then, if both sets are combined, the number of data is increased while the number of unknowns remains constant. Hence, the total data may now be enough for the complete set of equations to be soluble.

To make the unknowns the same, the ordinary and extraordinary virtual height curves must be sampled at frequencies which correspond to the same true heights of reflection. Considering one such pair of frequencies, this is to say that the values of f_R must be the same for the two modes. From the expressions for f_R ((5) and (7)), this requirement is seen to be met if the sample frequency f_n for the ordinary mode is made equal to $\tilde{f}_n$ for the extraordinary; the corresponding sample frequency f_n for the extraordinary may be found from (11). Thus the corresponding values of f_n are not equal for the two modes; moreover, if the sample frequencies are spaced evenly for the ordinary mode, then those for the extraordinary mode will be spaced unevenly. From here on, the two sets of sample frequencies will be assumed to be related in this way.

The question now arises, under what condition can the layer profile be determined fully by joint use of both virtual height curves? Suppose, to simplify the discussion, that all the data come in pairs; sample values of ordinary virtual height for which no corresponding extraordinary virtual heights are measurable, or vice versa, will be ignored. Suppose also that the range of index n for the two sets of data is $r \leqq n \leqq r+s-1$. Then the data comprise s values of h'_n, together with s values of $\tilde{h}_n$, a total of $2s$ separate items. The number of the unknowns Δz_m is $r+s$. Evidently the condition[4] under which the data determine the unknowns fully is that $s \geqq r$.

If the positions of the pairs of sample frequencies can be chosen arbitrarily, it is easy to arrange for a surplus of data. If, on the other hand, it is decided to space the frequencies evenly for one or other of the modes, say the ordinary, then the condition entails that the frequency range over which the ordinary virtual height is measurable should be wider than the unobserved range of low frequencies. Since, for most ionosondes, the low frequency limit lies below 2 Mc, while the ordinary penetration frequency for the $F2$ layer usually exceeds 4 Mc, this requirement is likely to be met most of the time.

Assume, then, that there are enough data to determine the unknowns. Two distinct cases are to be considered: In the first, s is equal to r, so that the data determine the unknowns exactly; in the second, s is greater than r, and the unknowns are over-determined. The exactly-determined case is treated in section 3.2, and the overdetermined case in section 3.3.

[4] Strictly speaking, this condition is necessary but not sufficient, for, even when it is satisfied, the determinant of the set of equations still might vanish. In practice, the determinant is unlikely to vanish exactly, though it might become very small. Then quite small errors in the data would produce large errors in the inferred values of the unknowns, a state of affairs that is to be avoided. For a discussion of the effects of random errors, see section 3.5.

3.2. The Exactly-Determined Case

In the exactly-determined case, the index n for the data (h'_n and $\tilde{h}_n$) runs from r to $2r-1$. The matrix eqs (17) and (19) together represent a set of $2r$ simultaneous equations, which are to be solved for the $2r$ unknown values of the Δz_m.

It is convenient to split the solution into two steps: First, the equations are solved for the increments of height in the unobserved frequency range ($0 \leqq m < r$); after which, the remaining increments in the observed range ($r \leqq m \leq 2r-1$) are found by existing methods.

In finding the Δz_m for the range $0 \leqq m < r$, the values outside this range first must be eliminated. This is done most simply by starting from the inverse eqs (21) and (22). Because the sample frequencies for the ordinary and extraordinary virtual height curves correspond to the same true heights of reflection, these two expressions must give the same values for the Δz_n to the limits of accuracy of the matrix formulas. Therefore, they can be equated to yield the relation that holds between the two virtual height curves as a consequence of their being derived from the same layer profile:

$$\sum_{m=0}^{n} \{\tilde{N}_{nm}\tilde{h}_n - N_{nm}h'_m\} \simeq 0. \tag{25}$$

When $n \geqq r$, (25) provides a relation between the observed and the unobserved virtual heights. This fact can be made plain by splitting the sum into separate parts for the two frequency ranges:

$$\sum_{m=r}^{n} \{N_{nm}h'_m - \tilde{N}_{nm}\tilde{h}_m\} \simeq \sum_{m=0}^{r-1} \{\tilde{N}_{nm}\tilde{h}_m - N_{nm}h'_m\} \; (n \geq r). \tag{26}$$

The next step is to substitute for the unobserved virtual heights and reintroduce the corresponding Δz_m by use of (17) and (19):

$$\sum_{m=r}^{n} \{N_{nm}h'_m - \tilde{N}_{nm}\tilde{h}_m\} \simeq \sum_{m=0}^{r-1} \left\{ \tilde{N}_{nm} \left(\sum_{l=0}^{m} \tilde{M}_{ml} \Delta z_l \right) - N_{nm} \left(\sum_{l=0}^{m} M_{ml} \Delta z_l \right) \right\} \quad (n \geq r)$$

$$\simeq \sum_{m=0}^{r-1} \sum_{l=0}^{m} \{\tilde{N}_{nm}\tilde{M}_{ml} - N_{nm}M_{ml}\} \Delta z_l \quad (n \geq r) \tag{27}$$

$$\simeq \sum_{l=0}^{r-1} \sum_{m=l}^{r-1} \{\tilde{N}_{nm}\tilde{M}_{ml} - N_{nm}M_{ml}\} \Delta z_l \quad (n \geq r) \tag{28}$$

The reversal of the order of summation on the right-hand side, between (27) and (28), may be justified by pointing out that, in the form of (27) a given m occurs in combination with all the l's from 0 to m; hence, a given l must occur in combination with all the m's from l to $r-1$, as indicated in (28). Or again, it may be checked simply by writing down the triangular array of the possible combinations of l and m.

Now (28) may be written in the concise form

$$S_n \simeq \sum_{l=0}^{r-1} P_{nl} \Delta z_l \qquad (n \geq r) \tag{29}$$

where

$$S_n = \sum_{m=r}^{n} \{N_{nm}h'_m - \tilde{N}_{nm}\tilde{h}_m\} \qquad (n \geq r) \tag{30}$$

and

$$P_{nl} = \sum_{m=l}^{r-1} \{\tilde{N}_{nm}\tilde{M}_{ml} - N_{nm}M_{ml}\} \qquad (n \geq r). \tag{31}$$

The matrix eq (29) relates the Δz_l for the unobserved frequency range to the quantities S_n, which are formed from pairs of corresponding values of ordinary and extraordinary virtual height in the observed range. Its solution will be written as

$$\Delta z_l \simeq \sum_{n=r}^{2r-1} Q_{ln} S_n \qquad (l \leq r-1), \tag{32}$$

where the matrices $[Q_{ln}]$ and $[P_{nl}]$ are mutually inverse.

Having thus arrived at the height increments Δz_l for $l \leq r-1$, the remaining values can be found by solving one of the original matrix equations row-by-row (see sec. 2.4). Use of the virtual height data for either mode should lead to identical results, even when the data contain errors of measurement, because the equations for the two modes together form an exactly-determined set.

3.3. The Overdetermined Case

In the overdetermined case, where $s > r$, any subset of $r+s$ selections from the $2s$ items of data would suffice, in principle, to determine the unknowns completely. If the data were free from error, it would not matter which subset was chosen. But when the data contain errors of measurement, some subsets will give better estimates than others. A still better estimate might be obtained by taking an appropriately weighted mean of the estimates from all the possible subsets, in which case each unknown would be given as a linear combination of all the data. The choice of the best combination would be governed by the statistics of the errors. Here it will be assumed, for convenience, that the errors of measurement of the virtual heights are all random, mutually uncorrelated, and distributed normally with equal standard deviations. Then the best estimate of the Δz_l is that given by the "Method of Least Squares" (see, for instance, Whittaker and Robinson [8]), and it is the solution of the matrix equation

$$R_m \simeq \sum_{l=0}^{r+s-1} U_{ml} \Delta z_l, \tag{33}$$

where

$$R_m = \sum_{n=r}^{r+s-1} \{M_{nm} h'_n + \tilde{M}_{nm} \tilde{h}_n\} \tag{34}$$

and $[U_{ml}]$ is a square symmetric matrix, the elements of which are given by the expression

$$U_{ml} = \sum_{n=r}^{r+s-1} \{M_{nm} M_{nl} + \tilde{M}_{nm} \tilde{M}_{nl}\}. \tag{35}$$

These equations are derived in appendix B.

The matrix eq (33) represents a new set of $r+s$ linear simultaneous equations, equal in number to the unknowns; each of the new equations is a linear combination of the $2s$ original equations, represented by (17) and (19). The solution of the complete set will be written, in matrix form, as

$$\Delta z_l \simeq \sum_{m=0}^{r+s-1} V_{lm} R_m, \tag{36}$$

where the matrix $[V_{lm}]$ is the inverse of $[U_{lm}]$ and likewise is square symmetric.

In eqs (33) to (36), the index r would be a constant, fixed by the properties of the recording equipment, but the index s would vary from one set of data to another because of the variations in the penetration frequencies of the $F2$ layer. Equation (36) is the solution of (33) just so long as $s \geq r$. The equations do not contain this restriction explicitly, but when $s < r$ the determinant of $[U_{ml}]$ vanishes so that the matrix cannot be inverted. Equation (36) is indeed applicable to the exactly-determined case when $s=r$, but for this special case the equations developed in section 3.2 give the same answer more simply; at least, the matrix that has to be inverted is smaller.

3.4. Some Practical Considerations

The variability of the index s makes the full least-squares solution inconvenient for practical use because each value of s requires a different matrix V_{lm} to be computed. In practice, it is best to fix s at some particular value, so that V_{lm} need be computed only once. The chosen value of s should be sufficiently high for the unknowns to be thoroughly overdetermined, yet sufficiently low for the uppermost sample frequency (f_{r+s-1}) never to exceed f_0F_2. The use of this particular matrix $[V_{lm}]$ yields the values of the Δz_l for $0 \leqq l \leqq r+s-1$. When data are available at frequencies above f_{r+s-1}, the remaining Δz_l may be obtained by continuing the row-by-row solution upwards. If the data contain random errors, however, the results would be different according as the ordinary or extraordinary virtual heights were used in the solution. Better results would be obtained by combining both sets of virtual height data following the Method of Least Squares. The least-squares value for each of these $\Delta z_l (l \geq r+s)$ is found by applying the procedure of section 2.4 to the equation

$$\Delta z_l \simeq \frac{M_{ll}}{M_{ll}^2 + \tilde{M}_{ll}^2} \left\{ h'_l - \sum_{m=0}^{l-1} M_{lm} \Delta z_m \right\} + \frac{\tilde{M}_{ll}}{M_{ll}^2 + \tilde{M}_{ll}^2} \left\{ \tilde{h}_l - \sum_{m=0}^{l-1} \tilde{M}_{lm} \Delta z_m \right\} \tag{37}$$

which replaces (24).

Probably not much accuracy is lost by fixing the value of s in the way suggested, which entails that only data at frequencies between f_r and f_{r+s-1} are used to determine the distribution of the low-density ionization. For this ionization has its greatest effects on the observed virtual heights at frequencies just above f_r, and hence these frequencies must provide the data that are most useful in determining its distribution.

Furthermore, the virtual height data at the higher frequencies may be systematically unreliable as a result of "lateral deviation." This effect is discussed more fully in section 5. Here it is mentioned only as a further justification for omitting the data at frequencies above f_{r+s-1} initially, when using the matrix method for the overdetermined case.

3.5. The Effect of Random Errors

If the measurements of the virtual heights contain random errors, these errors produce fluctuations in the inferred values of the Δz_l. The standard deviation of these fluctuations will be calculated now for the total range of l. The calculation will be made for the overdetermined case assuming that the Δz_l have been obtained by the full least-squares procedure ((33) to (36)). As before, the errors in the measurements will all be assumed to have the same standard deviation, and this will be called σ.

Now the change that occurs in Δz_l, as a result of a small error $\delta h'_n$ in the measurement of the ordinary virtual height h'_n, is found by differentiating (36) and (34):

$$\delta(\Delta z_l) \simeq \sum_{m=0}^{r+s-1} V_{lm} \left(\frac{\partial R_m}{\partial h'_n} \right) \delta h'_n \simeq \sum_{m=0}^{r+s-1} V_{lm} M_{nm} \delta h'_n. \tag{38}$$

An error in the measurement of the extraordinary virtual height $\tilde{h}_n$ has a similar effect. The total standard deviation of Δz_l, due to all the errors, is found by squaring and adding these two contributions, summing over all n, averaging, and then taking the square root. The results can be written as:

$$\text{Standard deviation of } \Delta z_l = \sigma \rho_l, \tag{39}$$

where

$$\rho_l^2 = \sum_{n=r}^{r+s-1} \left\{ \left[\sum_{m=0}^{r+s-1} V_{lm} M_{nm} \right]^2 + \left[\sum_{m=0}^{r+s-1} V_{lm} \tilde{M}_{nm} \right]^2 \right\}. \tag{40}$$

The "error coefficient" ρ_l, which is the standard deviation of Δz_l per unit standard deviation of the original measurements, is a measure of the susceptibility of the estimate of Δz_l to disturbance by random errors of measurement; the smaller is ρ_l, the more reliable is the estimate.

For the exactly-determined case, the error coefficients may be obtained by using (40) with s set equal to r. The special equations developed for this case in section 3.2 do not lend themselves well to the derivation of the coefficients.

Let it be emphasized that the ρ_l are the error coefficients of the *increments* of height (Δz_l) on the layer profile. Of greater interest, perhaps, are the error coefficients of the heights (z_l) themselves. These coefficients will be called π_l, and they are defined thus, by analogy with the ρ_l:

$$\text{Standard deviation of } z_l = \sigma \pi_l. \tag{41}$$

The π_l cannot be derived from the ρ_l in any simple way because the errors in the calculated Δz_l are mutually correlated even though there is supposed to be no correlation between the errors in the virtual height data. An expression may be derived for the π_l, however, similarly to that for the ρ_l, using (23) as well as (36) and (34); it is

$$\pi_l^2 = \sum_{n=r}^{r+s-1} \left\{ \left[\sum_{k=0}^{l} \sum_{m=0}^{r+s-1} V_{km} M_{nm} \right]^2 + \left[\sum_{k=0}^{l} \sum_{m=0}^{r+s-1} V_{km} \tilde{M}_{rm} \right]^2 \right\}. \tag{42}$$

Alternatively, in appendix C it is shown that the whole analysis can be developed from the start in terms of the z_l rather than the Δz_l, without altering the form of most of the expressions. In the analysis so developed, the π_l would be given by the equation analogous to (40).

4. A Test of the Method

The foregoing analysis has been tested on an artificial ionogram. The test consisted of assuming a particular layer profile, computing the corresponding ionogram (fig. 2), and then attempting to recover the profile from the data in the ionogram, using first the conventional matrix method, and then the various extensions of the method that were developed in section 3. In the numerical calculations, the gyrofrequency was taken as 1450 kc, and the angle of dip as 68.2°, which are the values that apply at Boulder, Colo. The calculations were performed on the IBM–650 computer at the Central Radio Propagation Laboratory of NBS.

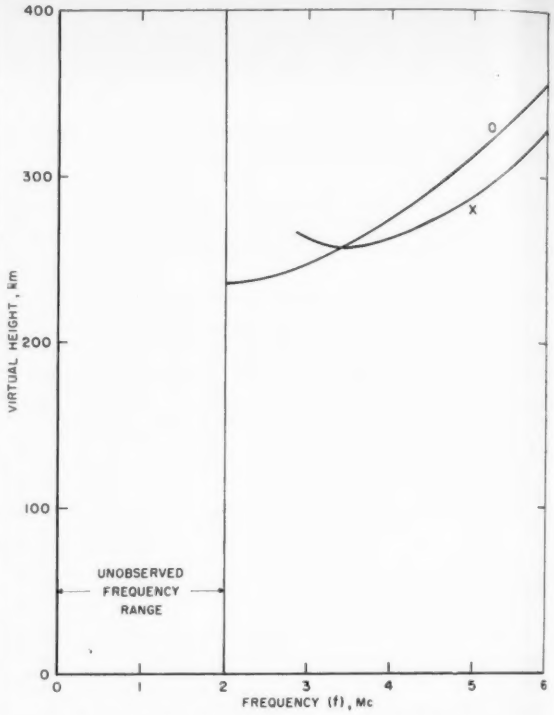


FIGURE 2. *The virtual height curves calculated for the assumed layer profile.*

The assumed layer profile is graphed as the solid curve in figures 3 to 5. It is monotonic, since the analysis is restricted to profiles of this type for the time being. It represents a distribution of ionization in which there is a ledge of low density beneath a layer of high density. Such a distribution might arise in the ionosphere soon after dusk, when the E-layer is decaying rapidly but the F-layer persists.

The virtual height curves that are derived from this profile are shown in the "ionogram" of figure 2. The lowest observable frequency (f_r) for the ordinary virtual height curve has been taken as 2 Mc, while the corresponding lower limit for the extraordinary curve is 2.85 Mc. The figure shows only those portions of the curves that are supposed to be observable. A particularly precise type of numerical integration (a 48-point Legendre-Gauss quadrature [9]) was used in computing the virtual heights, so that these data essentially are free from error. Hence any errors that may arise in the deduced layer profiles represent systematic defects in the methods of analysis.

The results of various attempts to recover the profile are presented in figures 3 to 5. In these figures, the points on the deduced profile are shown as solid dots, while the broken line is a direct interpolation between adjacent points.

Figure 3 shows the results of analyzing the ionogram by the conventional matrix method in which only the ordinary virtual height data are used [1, 3]. Because the low-density ionization has been ignored, all the computed points are much in error, lying well above the true profile (the solid curve); the error amounts to more than 30 km at the lowest observed frequency (2 Mc), while even at 6 Mc it is about 6 km. It is this error that the extensions of the matrix method seek to eliminate.

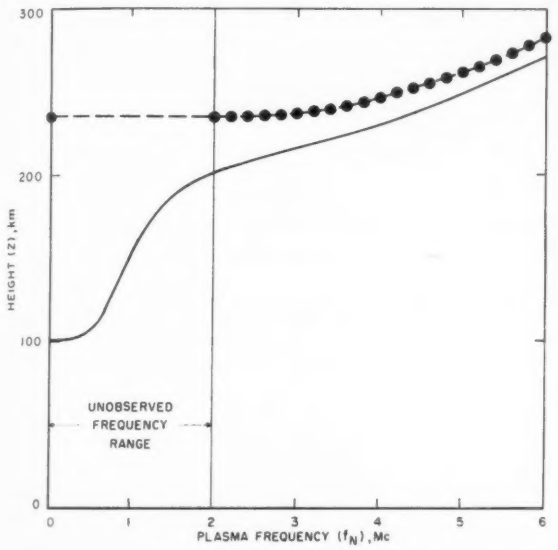


FIGURE 3. *The layer profile inferred from analysis of the ordinary virtual height curve by the conventional matrix method.*

For further explanation, see section 4 of the text.

However, in attempting to apply the new methods various problems were encountered, as follows:

The first, a comparatively trivial problem, was presented by the "rounding-off errors" in the digital computer. Such errors proved to be particularly serious in the inversion of the matrix $[U_{lm}]$. Eventually they were made negligible by performing the inversion with a "double-precision" program in which the computer carried 18 significant digits of each number instead of the usual 8 [10].

The main problem proved to be one of balance between the effects of random and systematic errors on the derived layer profile. The random errors are introduced by the inaccuracy of the virtual height data, as explained previously. The systematic errors are introduced when, in the process of converting the original integral equations into the corresponding matrix equations, the continuous curve of the profile is approximated by a series of rectilinear segments (see sec. 2.2.). The need to compromise between the two types of error does not arise in the analysis of a single virtual height curve, that is, in the situation studied in section 2, for in that situation the systematic errors can be made arbitrarily small by increasing the number of sample frequencies and reducing the intervals between them. The introduction of a new sample frequency adds a new unknown to the set of equations, but it also adds a new item of data that is related closely to this particular unknown, so that the random errors are relatively unaffected. The situation is quite different, however, when the virtual height curves are incomplete. Then the introduction of a new sample frequency in the unobserved range increases the number of unknowns without adding to the data, so that the precision with which the profile can be determined, in the presence of random errors, suffers accordingly. It is clear that the disposition of the sample frequencies, particularly of those in the unobserved frequency range, is a delicate problem and one that involves balancing the effects of random and systematic errors against each other. In view of the complexity of the equations, it seemed futile to attempt an analytic solution to this problem; instead, the sample frequencies were selected by a process of trial and error. The importance of the random errors was judged from the size of the error coefficients, and that of the systematic errors from the accuracy with which the derived profile reproduced the true one. An account of the series of trials is given now.

In the search for a satisfactory set of sample frequencies, five particular cases were tried in succession. Their properties are summarized in table 1, which gives for each case the values of the parameters r and s, the sample frequency f_l that corresponds to each index l, and the error coefficient ρ_l of each height increment Δz_l on the layer profiles. The profiles themselves were computed only for those cases where the error coefficients were reasonably small (i.e., <10), since it was found that the other cases developed severe "rounding-off errors" in the matrix multiplication (36).

Cases 1 and 2

In the first two trials a constant interval of 400 kc was maintained between sample frequencies. Case 1 was exactly-determined, with the number of data equal to the number of unknowns, while case 2 was overdetermined, with twice as many data as unknowns. As would be expected, the error coefficients for the overdetermined case were smaller than those for the exactly-determined case. In both cases, however, the error coefficients for the height increments in the observed frequency range were excessively large, of the order of several hundred. This fact suggests that it is not practical to space all the sample frequencies uniformly because random errors then have too severe an effect on the lower part of the deduced profile. Rather, a relatively coarse spacing must be employed in the unobserved frequency range, so that in this range the number of sample frequencies is reduced and the unknowns become very much overdetermined.

SAMPLE FREQUENCY	CASE 1		CASE 2		CASE 3		CASE 4		CASE 5	
	r = 5	s = 5	r = 5	s = 10	r = 1	s = 10	r = 3	s = 12	r = 3	s = 12
f_ℓ (Mc/s)	ℓ	ρ_ℓ	ℓ	ρ_ℓ	ℓ	ρ_ℓ	ℓ	ρ_ℓ	ℓ	ρ_ℓ
0.0	0	1.9 (2)	0	1.2 (2)	0	5.1 (0)	0	6.2 (2)		
0.2										
0.4	1	4.5 (1)	1	3.7 (2)						
0.6										
0.8	2	1.1 (2)	2	1.9 (2)			1	7.0 (2)	1	3.1 (0)
1.0										
1.2	3	2.5 (2)	3	1.6 (2)						
1.4										
1.6	4	1.1 (2)	4	7.8 (1)			2	7.2 (1)	2	4.0 (0)
1.8										
2.0	5	5.3 (0)	5	3.4 (0)	1	2.7 (0)	3	2.5 (0)	3	1.2 (0)
2.2							4	1.1 (0)	4	3.9 (-1)
2.4	6	3.9 (0)	6	2.6 (0)	2	4.1 (-1)	5	8.8 (-1)	5	3.0 (-1)
2.6							6	7.1 (-1)	6	2.4 (-1)
2.8	7	2.6 (0)	7	1.7 (0)	3	3.1 (-1)	7	5.8 (-1)	7	2.0 (-1)
3.0							8	4.8 (-1)	8	1.8 (-1)
3.2	8	1.8 (0)	8	1.2 (0)	4	2.5 (-1)	9	4.0 (-1)	9	1.6 (-1)
3.4							10	3.4 (-1)	10	1.4 (-1)
3.6	9	1.3 (0)	9	8.6 (-1)	5	2.1 (-1)	11	3.0 (-1)	11	1.3 (-1)
3.8							12	2.6 (-1)	12	1.2 (-1)
4.0			10	6.5 (-1)	6	1.8 (-1)	13	2.3 (-1)	13	1.1 (-1)
4.2							14	2.1 (-1)	14	1.1 (-1)
4.4			11	5.1 (-1)	7	1.6 (-1)				
4.6										
4.8			12	4.1 (-1)	8	1.5 (-1)				
5.0										
5.2			13	3.4 (-1)	9	1.4 (-1)				
5.4										
5.6			14	2.9 (-1)	10	1.3 (-1)				

TABLE 1. *The error coefficients (ρ_l) of the increments of height on the layer profile.*

Each coefficient is given in the form of a number, followed in parentheses by the power of ten by which that number should be multiplied; these coefficients are dimensionless. For further explanation, see section 4 of the text.

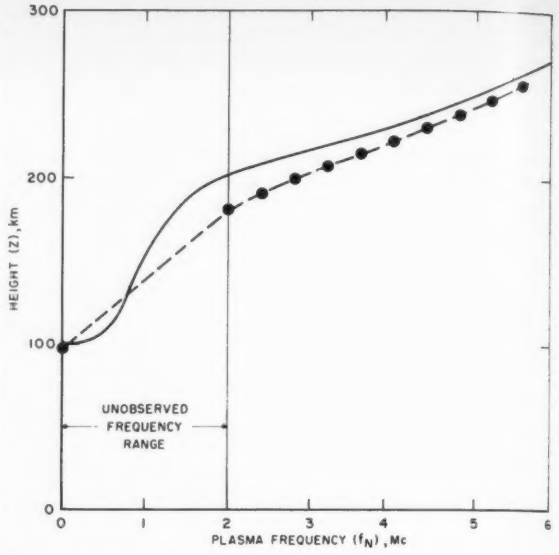


FIGURE 4. *The layer profile inferred from both virtual height curves by the extended matrix method in case 3.*

For further explanation, see section 4 of the text.

Case 3

Accordingly, in case 3 the effect was tried of representing the entire unobserved frequency range by a single interval of sample frequency; in other words, the variation of height with plasma frequency was assumed to be linear from 0 to 2 Mc. The sample frequencies in the unobserved range, however, were the same as those in case 2. The error coefficients now became reasonably small, and so the layer profile was computed for this case; the result is shown in figure 4. Although this profile is slightly more accurate than that of figure 3, it still contains large systematic errors which are due to the crudity of the straight-line approximation to the layer profile in the unobserved frequency range. Evidently a representation has to be found for this part of the profile that is intermediate in complexity between those of cases 2 and 3.

Case 4

In case 4, therefore, three sample frequencies were placed in the unobserved range, as against five for case 2, and one for case 3. Also, the number of sample frequencies in the observed range was increased, and their spacing was reduced to 200 kc in the hope that these changes would help to reduce the error coefficients. In fact, however, the error coefficients for index 0 and 1 turned out to be disappointingly large. The conclusion to be drawn from this result is that no reliable detail can be obtained at the low-frequency end of the unobserved range.

Case 5

In case 5, the final trial, the sample frequencies were the same as those in case 4, but the ionization in the lowest interval of sample frequency was ignored: from zero frequency up to f_1 (0.8 Mc) the layer profile was assumed to be horizontal. This restriction was imposed on the profile by modifying the matrices $[M_{nm}]$ and $[\tilde{M}_{nm}]$. The modifications consisted of omitting the row and column with index zero, and then defining the elements of the first column to be equal to unity. These changes had the desired effect of reducing the error coefficients to acceptable values, while at the same time bringing the inferred layer profile (fig. 5) into close agreement with the true profile. The only outstanding discrepancy is at the extreme low-frequency end of the unobserved range, where the assumption that the profile is horizontal makes it inevitable. Throughout the observed frequency range, the systematic errors of the computed points are all less than 1 km. Hence, in this final case, the choice of sample frequencies is deemed satisfactory.

No doubt the method could be improved further by continuing this process of trial and error, but the results already achieved are felt to give a sufficient demonstration of its potentialities.

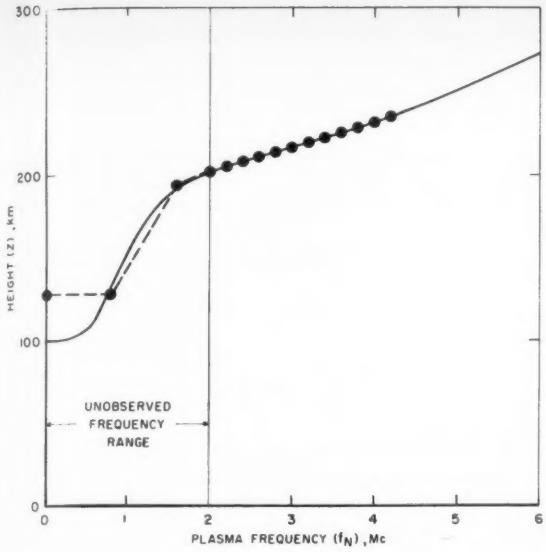


FIGURE 5. *The layer profile inferred by the extended matrix method in case 5.*

For further explanation, see section 4 of the text.

5. Discussion

This section contains a discussion of one difficulty that might be encountered in the practical use of the methods of this paper. The difficulty, noted previously in section 3.4, is that caused by "lateral deviation." It is well known that, in vertical sounding, the paths of the ordinary and extraordinary rays within the ionosphere are deviated laterally in opposite directions [11]. So long as the ionosphere is uniform in the horizontal plane, this deviation has no effect on the virtual heights. But whenever there are horizontal gradients of electron density, then the ordinary and extraordinary rays are propagated through different distributions of ionization, and for neither mode does the virtual height curve correspond to the vertical distribution of ionization directly above the sounder. This effect has been detected already as a source of error in attempts to measure the gyrofrequency in the F-layer by comparison of the penetration frequencies for the two modes [11]. Presumably it also would cause errors in attempts to deduce the vertical distribution of ionization by joint use of both virtual height curves.

In this connection, it is of interest that the comparison of the penetration frequencies observed on the ionograms indicates whether or not a horizontal gradient of electron density is present. If this comparison yields the correct value for the gyrofrequency, then there can be no horizontal gradient; such an ionogram could be analyzed with confidence to obtain the vertical distribution of ionization. If, on the other hand, there is a discrepancy between the calculated and true gyrofrequencies, then the size of this discrepancy provides a measure of the gradient; perhaps this information could be used to apply some first-order correction to the virtual height curves, so as to make them more amenable to analysis.

6. Conclusions

In principle, given an ionogram in which the virtual heights are not recorded below some lower limiting frequency, it is possible to deduce the complete layer profile for the full range of frequency by combinining the data in the ordinary and extraordinary virtual height curves; it must be presupposed, however, that the profile is monotonic and that the data are exact.

In practice the data contain random errors of measurement, and these errors have a severe effect on the inferred profile if too much information is sought about its shape in the unobserved frequency range. To keep the effect of random errors within reasonable bounds, detail must be sacrificed in the unobserved frequency range, particularly at the lowest frequencies; this is to say that the lower part of the profile must be characterized by a few parameters only, and that these parameters must be grossly overdetermined by the data. The abandonment of this low-frequency detail involves systematic errors even when the data are exact. In fact, a balance has to be struck between random and systematic errors. It is possible to strike a balance that gives a considerable improvement over the conventional matrix method, in which the ordinary virtual-height curve alone is analyzed.

The methods need further extension to cover situations where the layer profile is multivalued.

The author is indebted to T. E. Van Zandt, E. L. Crow, S. Chapman, J. W. Wright, and J. M. Watts for discussion and advice, and to Mrs. G. A. Hessing, J. W. Finney, and R. B. Norton for assistance with the numerical computations.

7. Appendixes

7.1. Appendix A: Formulas for Computing the Group Refractive Index Matrices

The elements of the matrices $[M_{nm}]$ and $[\tilde{M}_{nm}]$ for $1 \leqq m \leqq n$, are given by the integrals in (18) and (20), respectively. These integrals cannot be evaluated analytically, in general, so they have to be evaluated numerically. As numerical integrations, however, they are difficult to perform for elements on or near the diagonal of the matrix ($m \simeq n$), since the plasma frequency f_N then ranges close to the critical value for reflection, where the group refractive index becomes infinite. Nonetheless, the elements themselves are all finite; and, in fact, each integral may be converted into a form that is usable right up to the diagonal by a simple change of variable [1].

The procedures for the ordinary and extraordinary matrices are slightly different, so they will be treated separately.

a. Ordinary Matrix

For the ordinary mode, the matrix elements are given by the integral

$$M_{nm}=\frac{1}{\Delta f_m}\int_{f_{m-1}}^{f_m}\mu'(f_n,f_N)df_N \qquad (1\leqq m\leqq n). \quad (43)$$

To obtain a form in which the integrand stays finite as f_N approaches f_n, the variable of integration is changed to

$$\phi=\sin^{-1}(f_N/f_n). \quad (44)$$

Then

$$M_{nm}=\frac{f_n}{\Delta f_m}\int_{\phi_{m-1}}^{\phi_m}\{(1-X_n)^{\frac{1}{2}}\mu'(f_n,f_n\sin\phi)\}d\phi \qquad (1\leqq m\leqq n), \quad (45)$$

where ϕ_m, Φ_{m-1} are the values of ϕ for which $f_N=f_m$, f_{m-1}, and where

$$X_n=(f_N/f_n)^2=\sin^2\phi. \quad (46)$$

If the frequency interval Δf_m is a constant, say Δf, so that $f_n=n\,\Delta f$, then

$$M_{nm}=n\int_{\phi_{m-1}}^{\phi_m}\{(1-X_n)^{\frac{1}{2}}\mu'(f_n,f_n\sin\phi)\}d\phi \qquad (1\leqq m\leqq n). \quad (47)$$

In evaluating this integral, the integrand is computed from the equations of the magneto-ionic theory, using the familiar variables

$$X=(f_N/f)^2 \quad (48)$$

$$Y=f_H/f \qquad Y_L=Y\cos\theta \qquad Y_T=Y\sin\theta \quad (49)$$

in which θ is the angle between the direction of propagation (the vertical) and the earth's magnetic field; the effect of collisions is ignored, as is usual in this type of work. From these variables, the following parameters are defined:[5]

$$S=\{Y_T^4+4Y_L^2(1-X)^2\}^{\frac{1}{2}}, \quad (50)$$

$$D=2(1-X)-E, \quad (51)$$

$$E=Y_T^2+S, \quad (52)$$

$$F=E+2Y_L^2(1-X). \quad (53)$$

[5] These parameters are similar to a set introduced by Shinn [12] who is responsible for several of the formulas quoted here. However, some of Shinn's definitions have been altered slightly.

It is convenient to compute the integrand as the product of two factors:

$$(1-X)^{\frac{1}{2}}\mu'=\{\mu\mu'\}\left\{\frac{1-X}{\mu^2}\right\}^{\frac{1}{2}}. \quad (54)$$

Here μ is the wave refractive index for the ordinary mode. Each of these factors, separately, also remains finite as f_N approaches f. They are given by

$$\mu\mu'=\mu^2+\frac{X}{F^2}\left\{E(E+2Y_L^2)-\frac{4Y_L^4(1-X)(1-X^2)}{S}\right\} \quad (55)$$

and

$$\frac{1-X}{\mu^2}=1-\frac{2Y_L^2X}{E+2Y_L^2}, \quad (56)$$

while, furthermore

$$\mu^2=1-\frac{EX}{F}. \quad (57)$$

b. Extraordinary Matrix

The integral for the elements of the extraordinary matrix is

$$\tilde{M}_{nm}=\frac{1}{\Delta\tilde{f}_m}\int_{\tilde{f}_{m-1}}^{\tilde{f}_m}\mu_x'(f_n,f_N)\,df_N \qquad (1\leqq m\leqq n) \quad (58)$$

where

$$f_n=\frac{1}{2}f_H[\{1+4(\tilde{f}_n/f_H)\}^{\frac{1}{2}}\pm 1] \quad (59)$$

and the sign is to be taken positive or negative according to whether f_n is greater or less than f_H. Equation (58) may be changed into a form suitable for numerical integration by making the substitution

$$\tilde{\phi}=\sin^{-1}(f_N/\tilde{f}_n). \quad (60)$$

Then

$$\tilde{M}_{nm}=\frac{\tilde{f}_n}{\Delta\tilde{f}_m}\int_{\tilde{\phi}_{m-1}}^{\tilde{\phi}_m}\{(1-\tilde{X}_n)^{\frac{1}{2}}\mu_x'(f_n,\tilde{f}_n\sin\tilde{\phi})\}d\tilde{\phi} \qquad (1\leqq m\leqq n) \quad (61)$$

where $\tilde{\phi}_m,\tilde{\phi}_{m-1}$ are the values of $\tilde{\phi}$ for which $f_N=\tilde{f}_m$, $\tilde{f}_{m-1}$, and where

$$\tilde{X}_n=(f_N/\tilde{f}_n)^2=\sin^2\tilde{\phi}. \quad (62)$$

For the case where the frequency interval $\Delta\tilde{f}_m$ is a constant, say $\Delta\tilde{f}$,

$$\tilde{M}_{nm}=n\int_{\tilde{\phi}_{m-1}}^{\tilde{\phi}_m}\{(1-\tilde{X}_n)^{\frac{1}{2}}\mu_x'(f_n,\tilde{f}_n\sin\tilde{\phi})\}d\tilde{\phi} \quad (1\leqq m\leqq n). \quad (63)$$

Again, the integrand is computed as the product of two factors. Here the factors are, firstly,

$$\mu_x\mu_x' = \mu_x^2 + \frac{2}{D}\left[1-\mu_x^2-X^2-\frac{(1-X^2)(1-\mu_x^2)Y_L^2}{S}\right] \quad (64)$$

and secondly, the square root of

$$\frac{1-\tilde{X}}{\mu_x^2} = \frac{D}{2(1\mp Y)(1-X\pm Y)}(1+2Y_L^2/E). \quad (65)$$

The wave refractive index for the extraordinary mode is given by

$$\mu_x^2 = 1-\frac{2X(1-X)}{D}. \quad (66)$$

Note also that

$$X = \tilde{X}(1\mp Y). \quad (67)$$

Here, and in (65) above, the upper sign applies to frequencies above f_H, and the lower to frequencies below f_H.

7.2. Appendix B: Treatment of the Overdetermined Case by the Method of Least Squares

The equations to be solved are

$$h_n' \simeq \sum_{l=0}^{n} M_{nl}\Delta z_l \qquad (r \leq n \leq r+s-1) \quad (17)$$

for the ordinary mode, and

$$\tilde{h}_n \simeq \sum_{l=0}^{n} \tilde{M}_{nl}\Delta z_l \qquad (r \leq n \leq r+s-1) \quad (19)$$

for the extraordinary. Observed values of the h_n' and $\tilde{h}_n$ are given for $r \leqq n \leqq r+s-1$, where $s \geqq r$, but these data contain random errors of measurement. The aim is to find a set of Δz_l which, if inserted in the equations, would yield calculated values for the h_n' and $\tilde{h}_n$ that agree with the observed values as closely as possible. The measure of disagreement is the sum of the squares of the differences between the two sets of values, and this quantity is to be minimized.

Now, if certain values are assumed for the Δz_l, the implied errors of measurement are

$$E_n = h_n' - \sum_{l=0}^{n} M_{nl}\Delta z_l \quad (68)$$

and

$$\tilde{E}_n = \tilde{h}_n - \sum_{l=0}^{n} \tilde{M}_{nl}\Delta z_l. \quad (69)$$

For the ordinary mode, the sum of the squares of the implied errors is

$$\sum_{n=r}^{r+s-1} E_n^2 = \sum_{n=r}^{r+s-1} (h_n')^2 - 2\sum_{n=r}^{r+s-1}\sum_{l=0}^{n} M_{nl}h_n'\Delta z_l + \sum_{n=r}^{r+s-1}\sum_{k=0}^{n}\sum_{l=0}^{n} M_{nk}M_{nl}\Delta z_k\Delta_l. \quad (70)$$

A similar expression applies to the extraordinary mode. These two expressions must be added, and the total minimized. The condition that it shall be a minimum with respect to variations of all the Δz_m is that

$$\frac{\partial}{\partial(\Delta z_m)} \sum_{n=r}^{r+s-1} \{E_n^2 + \tilde{E}_n^2\} = 0 \quad (71)$$

for all m $(0 \leqq m \leqq r+s-1)$. On evaluating the derivatives, this condition yields the Δz_l as the solutions of (33).

7.3. Appendix C: The Analysis Presented in Terms of the Ordinates z_l

In the analysis described in this paper, the ordinates of the layer profile have been arrived at by two distinct steps: First, various matrix equations have been set up and solved in terms of the height increments Δz_l, then these increments have been summed to obtain the actual ordinates z_l. It is possible to combine these two steps by rewriting the matrix equations with the ordinates themselves as the unknowns. Consider, for instance, the eq (17) for the ordinary virtual height. It may be rearranged as follows:

$$\begin{aligned} h_n' &\simeq \sum_{m=0}^{n} M_{nm}\Delta z_m \quad (17)\\ &\simeq M_{n0}\Delta z_0 + \sum_{m=1}^{n} M_{nm}\Delta z_m \\ &\simeq M_{n0}z_0 + \sum_{m=1}^{n} M_{nm}(z_m - z_{m-1}) \\ &\simeq M_{n0}z_0 + \sum_{m=1}^{n} M_{nm}z_m - \sum_{m=1}^{n} M_{nm}z_{m-1} \\ &\simeq \sum_{m=0}^{n} M_{nm}z_m - \sum_{m=0}^{n-1} M_{n,m+1}z_m \\ &\simeq \sum_{m=0}^{n} a_{nm}z_m, \quad (72) \end{aligned}$$

say,

where

$$a_{nm} = \begin{cases} M_{nm} - M_{n,m+1} & (m<n) \\ M_{nn} & (m=n) \\ 0 & (m>n). \end{cases} \quad (73)$$

This is the form in which Budden [1] uses the matrix equation; it is, of course, completely equivalent to the other.

The eq (19), for the extraordinary virtual height, may be rearranged in the same way:

$$\tilde{h}_n = \sum_{m=0}^{n} \tilde{a}_{nm} z_m, \qquad \text{say.} \tag{74}$$

Since the eqs (72) and (74) are of the same form as (17) and (19), it is clear that the analysis of this paper could have been based equally well on the new equations, and developed in terms of the ordinates z_l; the arguments would proceed just as before, with the z_l replacing the Δz_l, so there is no need to give the details. The choice of which of the two sets of equations to start from is largely a matter of convenience, though it may be that, in computation, one would involve smaller "rounding-off errors" than the other.

8. References

[1] K. G. Budden, A method for determining the variation of electron density with height ($N(z)$ curves) from curves of equivalent height against frequency ((h',f) curves), Rept. Cambridge Conf. Ionospheric Phys., Phys. Soc. London, p. 332 (1955).

[2] O. E. H. Rydbeck, A theoretical survey of the possibilities of determining the distribution of free electrons in the upper atmosphere, Trans. Chalmers Univ. Technol., Gothenburg, No. 3 (1942).

[3] E. R. Schmerling, The reduction of $h'(f)$ records to electron-density-height profiles, Penn. State Univ., Ionospheric Research Sci. Rept. No. 94 (1957).

[4] J. E. Jackson, A new method of obtaining the electron density profiles from $h'(f)$ records, J. Geophys. Research **61,** 107 (1956).

[5] J. O. Thomas, The distribution of electrons in the ionosphere, Proc. I.R.E. **47,** 162 (1959).

[6] J. E. Titheridge, The use of the extraordinary ray in the analysis of ionospheric records (in press).

[7] J. O. Thomas, J. Haselgrove, and A. Robbins, The electron distribution in the ionosphere over Slough—I. Quiet days, J. Atmospheric and Terres. Phys. **12,** 46 (1958).

[8] E. Whittaker and G. Robinson, The calculus of observations, (Blackie, London) 4th Ed., p. 209 (1944).

[9] F. B. Hildebrand, Introduction to numerical analysis, p. 323 (McGraw-Hill Book Co., Inc., New York, N.Y., 1956).

[10] J. D. Chappell, Double precision matrix inversion, IBM 650 Program Library, File No. 5.2.009.

[11] J. C. W. Scott, Longitudinal and transverse propagation in Canada, J. Geophys. Research **55,** 65 (1950).

[12] D. H. Shinn, Tables of group refractive index for the ordinary ray in the ionosphere, Rept. Cambridge Conf. Ionospheric Phys., Phys. Soc. London, p. 402 (1955).

Boulder, Colo.

(Paper 64D2–45)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March–April 1960

Measured Statistical Characteristics and Narrow-Band Teletype Message Errors on a Single-Sideband 600-Mile-Long Ultrahigh-Frequency Tropospheric Radio Link

E. F. Florman * and R. W. Plush *

(October 14, 1959)

Measurements of a 417 megacycles per second unmodulated radio carrier over a 600-mile tropospheric path indicated that the variations of the received carrier envelope amplitude with time over 30-minute periods roughly approximated a Rayleigh distribution in the majority of the tests. Cumulative time distributions of the carrier fade durations were obtained over a range of carrier envelope power levels and were found to resemble corresponding distributions for narrow band thermal noise. The fade rate of the carrier envelope, at the median power level, was less than 0.2 cycle per second; this comparatively low fade rate is thought to be due mainly to the narrow (1°) antenna beam widths that were used and the consequent low order of multipath propagation of the radio waves. The half hour median power levels of the received carrier varied over a range of approximately ±8 decibels. The effective low-pass bandwidth of the carrier envelope spectrum was found to vary from 0.06 to 0.17 cycle per second. The measured median transmission loss was approximately 183 decibels which is within 3.5 decibels of the calculated value for the summer afternoon hours covered by the tests.

With antennas spaced (normal to the path) at each end by 150 wavelengths it was found that parallel-path, divergent-path, or convergent-path types of transmission gave cross-correlation coefficients of the carrier envelopes which ranged from 0.08 to 0.20. For the same antenna spacing but using crossed-path type of transmission, the cross-correlation coefficient was 0.57. Diversity measurements of single-path crossed polarization type of transmission indicated that the cross-correlation coefficient of the carrier envelopes was very close to unity.

Nondiversity narrow-band FSK error measurements indicated that an 18-decibel signal-to-noise power ratio over an effective bandwidth of 290 cycles per second (at the limiter-discriminator input) for a fading signal resulted in 0.8 percent binary errors and 4.0 percent teletype character errors. Extrapolation of these results indicates that a signal-to-noise ratio of 27 decibels is required to reduce the teletype character error rate to 0.1 percent in the same effective bandwidth.

1. Introduction

The performance of a tropospheric beyond-the-horizon radio communication system involves the study of the dependence of tropospheric radio wave propagation characteristics on frequency, path length, antenna beam widths, angular distance, meteorological conditions, antenna spacing, etc. Experimental data on a variety of tropospheric paths appear to indicate the feasibility of considering the radio-frequency carrier as being roughly equivalent to random noise passing through a narrow band filter. The limits of the simulating system are being determined by experimental measurements on tropospheric radio links under conditions covering a wide range of frequencies, path lengths, and other pertinent factors.

The work covered by this report was carried out in order to obtain experimental data on the statistical behavior of tropospheric uhf radio wave propagation over a 600-mile path. This system represents to date the longest operating tropospheric radio link and also employs very narrow antenna beam widths of approximately 1 deg; hence, it represents extremes as regards both length of path and antenna beam width.

The information obtained on the tropospheric radio signals is to be applied in studies involving various types of modulation for the purpose of establishing engineering standards of performance.

2. Equipment and Technique

The tropospheric radio link used to make the experimental measurements was the "Domestic Dew Drop" system designed by Lincoln Laboratory personnel and constructed by the General Electric Co., with assistance in the installation work from the Page Engineering Co. The military designation of this system is AN/FRC–47.

A total of twenty-seven 30-min tests were made during the period from July 10 to 16, 1958, covering times of day ranging from 0900 to 2400 e.s.t. In addition, one test was made on May 22, 1958; further tests could not be made at that time due to the fact that the system was undergoing a series of adjustments. Horizontally polarized waves were received on 23 of the tests and a combination of horizontally polarized and vertically polarized waves were received on the remaining 5 tests. Two radio receiver

*Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

systems were used for tests Nos. 4 to 28 inclusive in order to obtain either space diversity data or to separate two signals which were being transmitted simultaneously over the same link at slightly different frequencies. Each radio receiving system was calibrated in terms of absolute input power levels before each test by means of a calibrated Micropot [1] as indicated in figure 1. Cross coupling between the

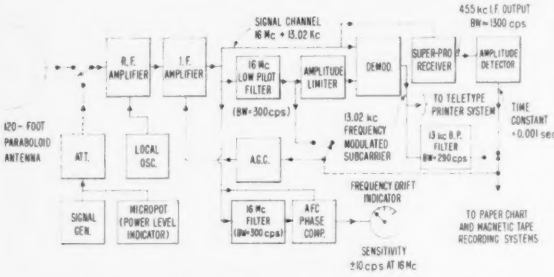


FIGURE 1. *Block diagram of receiving system.*

horizontally and vertically polarized receiving antennas was down 40 db. Radio transmission, at a carrier frequency of 417.05 Mc was from Millstone Hill, Mass., to Sauratown Mt., N.C.; the great circle path length being 638 miles.

Figure 1 is a block diagram of the radio receiving system used to carry out the tests. As indicated in the diagram, the system was arranged to be used either as a linear amplifier without AGC action or as a constant output-signal device with AGC action. When used as a linear system to obtain a measure of the propagation characteristics of the path, the received signal was taken from the 16-Mc IF output of the uhf tropospheric link radio receiver and amplified through a HF communications type of radio receiver. The 455-kc IF output of the communication receiver was then passed through an amplitude detector having an output circuit time constant of 0.001 sec and the resulting amplitude variations, which were representative of the amplitude variations of the carrier envelope, were recorded as frequency modulation of a subcarrier on magnetic tape. The output of the above amplitude detector was also recorded on paper in order to provide a visual check on the results. At the same time, amplitude distribution data on the carrier envelope were obtained by means of an array of counters each of which was adjusted to trigger at a different power level. A comparison of the amplitude distribution curves obtained directly at the receiving site with corresponding curves taken from the magnetic tape recording showed an overall discrepancy of less than ± 0.5 db.

During some of the tests involving teletype and binary message errors, power was transmitted simultaneously at frequencies of 417.05 and 417.06302 Mc by utilizing a sideband 13.02 kc above the carrier (low pilot) frequency. In this case, AGC was used and a FSK teletype or a binary signal was taken from the demodulator output of the General Electric receiver, passed through a 13.02-kc band pass filter (having an "effective" pass band of 290 cps), and then recorded directly on magnetic tape. This tape recording constituted a replica of the teletype (or binary) frequency modulation imposed on the 13.02-kc subcarrier at the radio transmitter; it includes propagation and radio receiver noise as well as carrier fading effects. The teletype signal consisted of the standard "quick brown fox" message which was also recorded by means of a teletype printer. The 10-cps binary signal (see fig. 10) was a voltage wave having a rise time of approximately 10 msec, and a flat-top duration of 30 milliseconds; this binary signal was generated by sending a continuous string of O's. A binary signal rate of 10 cps was the maximum frequency practically attainable by the system at the time the tests were made using only standard teletype equipment. For both the teletype and the binary signals, the frequency shift from "mark" to "space" was 60 cps. The envelope of the 417.05-Mc carrier was simultaneously recorded on a second receiving system (without AGC) as explained previously.

Carrier measurements involving the use of the receiver combining circuits are not included in this report because these circuits were not in proper adjustment at the time the system tests were made. AFC was not used during these tests; however, the frequency drift indicator showed that the drift in the local oscillator was never greater than ± 10 cps, which was within the 300-cps pass bands of the receiver system filters. Transmitting power levels were varied from approximately 500 w to 4 kw, being constant for any one test.

System data pertinent to the tests are tabulated as follows:

Millstone Hill, Mass., Station

1. Coordinates—42°36′48″ N; 71°29′38″ W.
2. Elevation—330 ft A.S.L. (ground level).
3. Azimuth of path, Millstone Hill to Sauratown Mt.—230°33′37″ clockwise from north.
4. Antennas—two, each 120 ft by 120 ft parabolic reflectors with feed horn at focus.
5. Estimated plane wave antenna gain—41.2 db. Antenna beam width approximately 1°.
6. Antenna spacing normal to propagation path—approximately 150 λ (at 417 Mc).
7. Antenna line loss—approximately 1 db.
8. Optical horizon—+15 min.
9. Distance to optical horizon—approximately 4 miles.
10. Distance from Millstone Hill to Sauratown Mt.—638 statute miles.

Sauratown Mt., N.C., Station

1. Coordinates—36°22′52″ N; 80° 21′ 22″ W
2. Elevation—1915 ft A.S.L. (ground level).
3. Azimuth of path, Sauratown Mt. to Millstone Hill—44°54′24″ clockwise from north.

[1] Furnished and calibrated by Section 1, Division 84, National Bureau of Standards, Boulder, Colo. This device was stated to have an absolute accuracy of ±3 percent.

4. Antennas—two, each 120 ft by 120 ft parabolic reflectors with feed horn at focus.
5. Estimated plane wave antenna gain—41.2 db. Antenna beam width approximately 1°.
6. Antenna spacing normal to propagation path—approximately 150 λ (at 417 Mc).
7. Antenna line loss—approximately 1 db.
8. Optical horizon—20 min.
9. Distance to optical horizon—>10 miles.
10. Angular distance[1] [2]—130.7 milliradians (calculated, corrected for refraction).
11. Receiver noise figure—approximately 5 db.
12. Measured equivalent noise power levels of the G.E. receivers over an "effective" bandwidth of approximately 150 cps—
    —148.0 dbm (decibels below 1 milliwatt) for receiver 1,
    —145.0 dbm for receiver 2,
    —146.5 dbm (estimated) for receiver 3, and
    —146.5 dbm (estimated) for receiver 4.

3. Experimental Results

The results of the tests are separated into two categories: (1) The effects of the path on the characteristics of the received radio carrier. These measurements were made using an unmodulated carrier at the low-pilot frequency of 417.05 Mc and (2) measurements of teletype and binary signal errors.

The factors involved in the first of the above two classifications of results are: power levels, amplitude distribution, fade rate (at median or other levels), fade duration distribution, power spectrum, variation of the 30-min medians, transmission loss, and diversity measurements; all of these factors were considered in terms of the envelope of the received radiofrequency carrier. The second category above deals with measurements of the teletype and binary signal errors in terms of carrier-to-noise (C/N) ratios and receiving system bandwidths.

3.1. Characteristics of the Received Carrier Envelope

The equipment used to obtain cumulative distributions of the received (and recorded) radio carrier envelope amplitudes and the fade rates was similar to that used by Watt and Maxwell [2]. The fade duration distributions were obtained by modifying the equipment so that only those fades exceeding a particular preset duration would be registered.

a. Received Power Levels and Amplitude Distributions

The range of power levels and the types of amplitude distributions obtained for the received carrier envelope are shown in figure 2. Preliminary test runs showed that sample times greater than 20 min were required in order to obtain a statistically reliable sample of the carrier envelope. In most

[1] Figures in brackets indicate the literature references at the end of this paper.

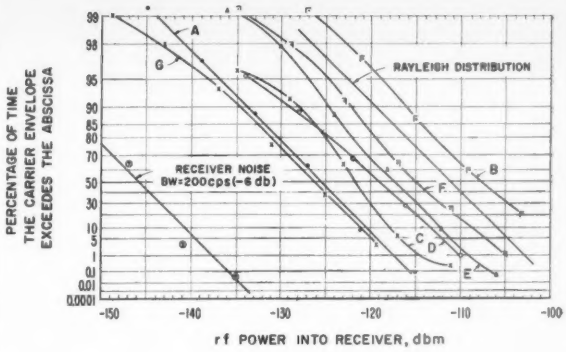


FIGURE 2. *Typical cumulative distributions of carrier envelope amplitudes.*

Sauratown Mt., N.C.
Carrier frequency=417.05 Mc
Transmission from Millstone Hill, Mass.
System BW dc to 42 cps and dc to 650 cps (−6db)
Transmitting and receiving antennas
120-ft square paraboloids
Sampling period 30 min

Run	Curve	Transmitter power	Date 1958	Time e.s.t.	Median	Transmission loss
		kw			*dbm*	*db*
1	A		5–22	1440	−125.6	
2	B	3	7–10	1745	−107.7	170.5
4	C	3	7–11	1350	−121.6	184.4
5	D	3	7–11	2020	−119.0	181.8
9	E	3	7–12	2030	−117.4	180.2
10	F	3	7–13	1300	−115.0	177.8
19	G	1	7–15	1450	−126.0	184.0

cases there is a similarity between the slopes of the measured distribution curves and the theoretical Rayleigh distribution. Curve C has the greatest departure from a Rayleigh distribution; however, the slope of a straight line best fitting the measured points closely approximates the slopes of the other measured distributions. As a matter of record, the data for curve C were obtained during a severe thunder storm at the receiving site and along most of the path of propagation, while data for curve D were obtained on the same day but 6 hr later when the electrical storm had subsided. The difference between curves C and D probably could be attributed to the effects of lightning discharges. The range of the 30-min median values of the received power over the period of six days is seen to be approximately 19 db; however, 5 db of this range can be attributed to a variation of the power level of the radiated carrier at the transmitter.

b. Fade Rates

The curves in figure 2 give an accurate statistical measure of the received power levels and the range of the variations of the received carrier envelope; however, these curves do not yield any information regarding the rates of fading, that is, the variation

of the carrier envelope with time, at various power levels. A more restricted definition of carrier fade rate is: the average number of times that the carrier envelope passes through its median power level in one direction, per second. Fade-rate curves measured for two typical cases at various power levels are shown in figures 3 and 4. These curves are contaminated by noise, particularly in the region of short duration fades (less than 100 msec where a fade duration is defined as the time spent below a specific level. A more accurate measure of the short duration fades of the carrier envelope alone would be obtained by extrapolating the "plateaus" of the curves into the short fade region. In previous work [3] this extrapolation procedure (for the median power level curves) yielded an approximate measure of the power spectrum width of the received carrier envelope. In this particular case we would obtain an average spectrum width of approximately 0.12 cps from the curves in figures 3 and 4. The pass band was limited to approximately 40 cps which was ample to pass most of the information on the carrier envelope amplitude variations, however, information on the noise characteristics above 40 cps that is, fade durations below 25 msec, would be inaccurate.

The received signals were also analyzed to determine the percentage of time that the carrier envelope faded below particular power levels for a fade duration equal to or greater than chosen values. Typical results of this analysis are shown in figures 5 and 6. A comparison between these sets of curves shows a greater percentage of short duration fade times for figure 5 than is shown in figure 6. The curves shown in figures 5 and 6 are useful for modulation studies where it is desired to estimate the percentage message errors, taking into account the carrier fade duration as well as the median carrier-to-noise ratio.

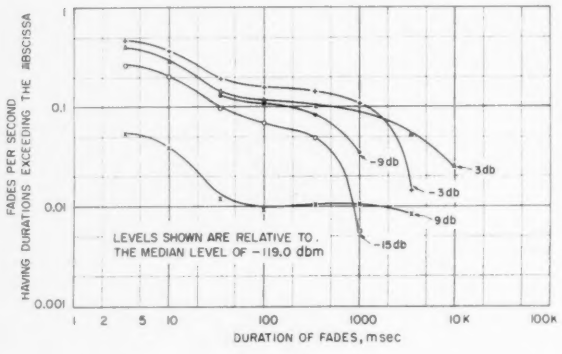


FIGURE 3. *Carrier envelope fade duration and rate at various power levels.*

Sauratown Mt., N.C.
Run 5, July 11, 1958—2008 to 2038 e.s.t.
Transmission from Millstone Hill, Mass.
Low pass bandwidth of system=42 cps (−6 db)
Carrier frequency=417.05 Mc
Transmitting and receiving antennas
120-ft square paraboloids

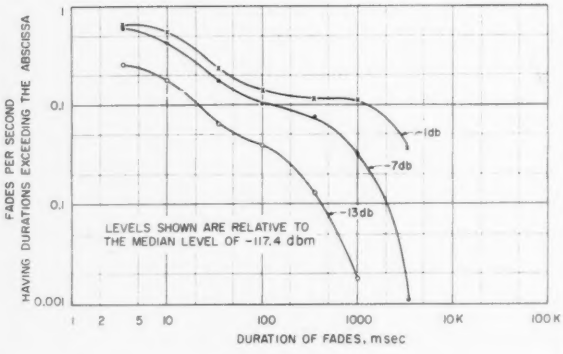


FIGURE 4. *Carrier envelope fade duration and rate at various power levels.*

Sauratown Mt., N.C.
Run 9, July 12, 1958—2016 to 2046 e.s.t.
Transmission from Millstone Hill, Mass.
Low pass bandwidth of system=42 cps (−6 db)
Carrier frequency=417.05 Mc
Transmitting and receiving antennas
120-ft square paraboloids

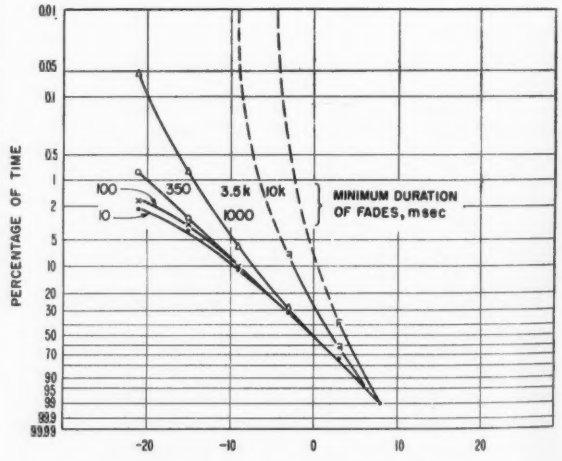


FIGURE 5. *Distribution of the carrier envelope fade duration showing the percentage of time that the carrier envelope fades to and remains below the level indicated by the abscissa, for a time duration exceeding the values indicated on the curves.*

Sauratown Mt., N.C.
Run 5, July 11, 1958—2008 to 2038 e.s.t.
Carrier frequency=417.05 Mc
Transmission from Millstone Hill, Mass.
Low pass bandwidth of system 42 cps (−6 db)
Transmitting and receiving antenna
120-ft square paraboloids
Median level=−119.0 dbm

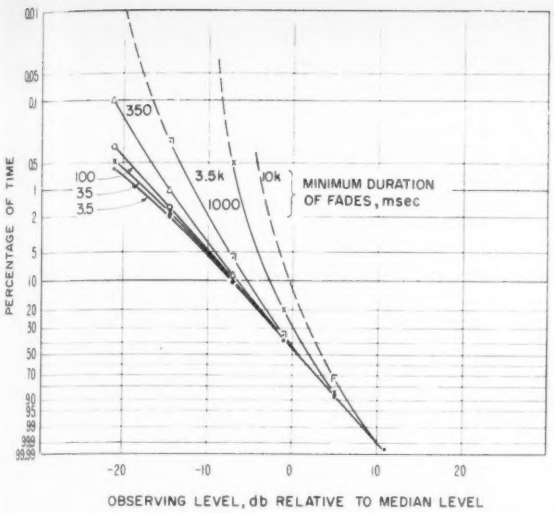


FIGURE 6. *Distribution of the carrier envelope fade duration showing the percentage of time that the carrier envelope fades to and remains below the level indicated by the abscissa, for a time duration exceeding the values indicated on the curves.*

Sauratown Mt., N.C.
Run 9, July 12, 1958—2016 to 2046 e.s.t.
Carrier frequency=417.05 Mc
Transmission from Millstone Hill, Mass.
Low pass bandwidth of system 42 cps (−6 db)
Transmitting and receiving antenna
120-ft square paraboloids
Median level=−117.4 dbm

c. Power Spectra

The power spectra of the carrier envelope, shown in figure 7, were obtained from playbacks of the magnetic tape recordings of the received carrier envelope. These results represent the bandwidth extremes for the tests covered by this report; selection was made in terms of extremes in fade rates. Discrete points on these curves were obtained by time averaging the power output level over an effective bandwidth of 0.041 cps at a sweep frequency over the spectrum, of 0.001 cps.

The effective pass band of the "propagation path noise" is seen to be considerably less (by a factor of more than 25 to 1) than the bandwidth measured on a shorter tropospheric path [3]. The effective bandwidths, B_{eff}, tabulated in figure 7 are defined as the widths of rectangular spectra having the same total power as the measured spectra and with constant power densities equal to the maximum power density of the measured spectra. The slower fade rate (or narrower propagation path noise pass band) for the longer path might be attributed to the following factors: (1) A difference in the average heights of the paths above the earth and a consequent difference in the average characteristics of the paths; (2) very narrow antenna patterns were used for the long path, which would result in a reduction of multipath effects; and (3) the 2-to-1 ratio of carrier frequencies would also account for some of the difference in fade rate over the two paths.

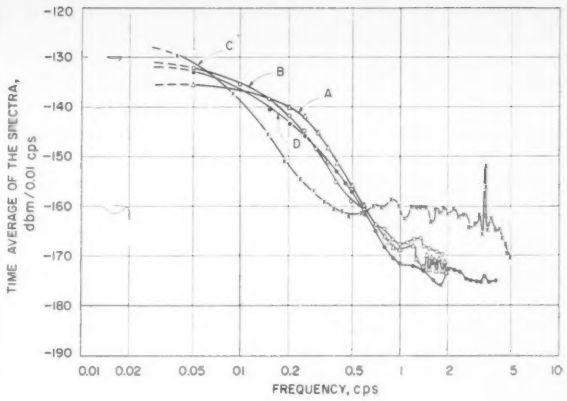


FIGURE 7. *Typical power spectra.*

Sauratown Mt., N.C.
Carrier frequency=417.05 Mc
Transmission from Millstone Hill, Mass.
System (low pass) bandwidth 7.1 cps (−6 db)
Transmitting and receiving antennas
120-ft square paraboloids
Sampling period, 30 min
July, 1958

Run	Curve	Date 1958	Time e.s.t.	Median power level of received signal	Total average spectrum power	B_{eff} of spectrum
				dbm	*dbm*	*cps*
5	A	7–11	2029	−119.0	−123.0	0.17
9	B	7–12	2030	−117.0	−120.4	0.06
10	C	7–13	1300	−114.5	−120.3	0.07
24	D	7–16	1315	.....	−121.9	0.06

The effect of reflections from airplanes [4] (along the path) is clearly evident from curve C in figure 7 where there is an enhancement of power at frequencies above 0.5 cps. In some instances it was noted that the carrier envelope faded at a very uniform and regular rate of up to 7 cps due apparently to reflections from airplanes. Lesser and varying amounts of "airplane fading" effects are seen from curves A, B, and D. It was also noted that lightning discharges have the effect of increasing the bandwidth of the propagation noise; this effect is perhaps due to a combination of reception of the energy from the lightning stroke and a short time enhancement of the carrier due to reflections from the ionized paths set up by the strokes.

It seems unlikely that a consideration only of these spectrum widths and spectrum power levels of the propagation noise would be sufficient to determine the effect of this noise on the message errors; due consideration should also be given to the duration of the carrier fades relative to the noise peaks at particular levels. However, the spectra shown by figure 7 do give a measure of the noise with which the signal must compete in terms of a resultant carrier amplitude or carrier envelope. The frequency modulation imposed on the resultant received carrier by the propagation path is expected to have a spectrum bandwidth of the same order as the above

spectrum of the path propagation amplitude modulation noise [4]. From the table in figure 7 it is seen that the total average power of the transmission-path noise lies below the median power levels from 3.4 to 5.8 db; for a Rayleigh distributed signal the total spectrum power (i.e., variance) should be 5 db below the median power level of the signal.

3.2. Transmission Loss

The predicted hourly median transmission loss distribution for the Millstone Hill of Sauratown Mt. tropospheric path is shown in figure 8; the probable error of the predicted median is plotted relative to the predicted median curve. This calculated distribution of the median transmission loss was made assuming that each antenna line loss was 1 db and that the antenna-to-medium coupling loss was 4.7 db for each antenna. The plane wave gain was estimated to be 73 db. If the assumption is made that the antennas are loss-free, that is, if the ohmic losses, ground losses, etc., are negligible, then the basic transmission loss or, what is also known as the path transmission loss, can be obtained by adding 73 db to the curve values in figure 8.

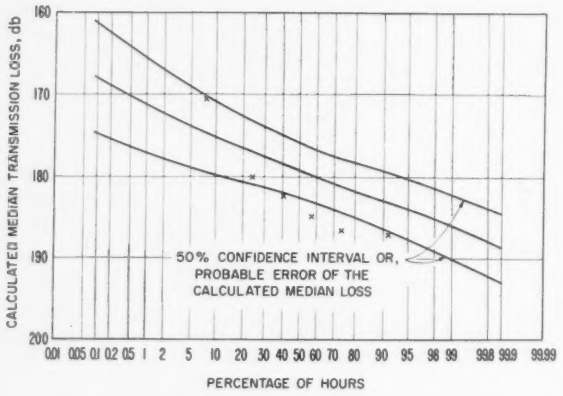

CALCULATED MEDIAN TRANSMISSION LOSS, db
160
170
180
190
200
50% CONFIDENCE INTERVAL OR,
PROBABLE ERROR OF THE
CALCULATED MEDIAN LOSS
0.01 0.05 0.1 0.2 0.5 1 2 5 10 20 30 40 50 60 70 80 90 95 98 99 99.8 99.9 99.99
PERCENTAGE OF HOURS

FIGURE 8. *Calculated transmission loss.*

Calculated distribution of hourly median of transmission loss for summer afternoon and evening
Transmission from Millstone Hill, Mass. to Sauratown Mt., N.C. frequency, 417 Mc
X-measured values, six half hour medians
Transmission and receiving antennas
120-ft square paraboloids
Path length, 638 statute mi
Angular distance, 130.7 milli-radians
Path antenna gain 73 db

Measurements of the transmission loss were made in terms of the power input to the transmitting terminals and the available median power level at the receiving antenna; an allowance of 1 db was made for antenna line loss at the transmitter and at the receiver. If the antennas are assumed to be loss-free, the above measurements of the system loss from the input to the transmitting antenna to the receiving antenna are equivalent to the transmission loss.

Measured values of the median transmission loss are tabulated in figure 2 and the cumulative distribution of these measured medians is shown in figure 8. The measured values of the median transmission loss are slightly higher than the predicted values, perhaps due to the limited number of samples. However, it should be noted that the estimated antenna to medium coupling losses and the possible improper pointing of the antennas are involved in the slight discrepancy between the predicted and the measured transmission loss.

3.3. Diversity Measurements

Five types of dual-diversity measurements were made; four of the measurements were made with spaced antennas and one was made using crossed polarization transmission with single antennas at each end of the radio link. The results of these tests are shown in table 1.

Figure 9 shows the antenna arrangement, polarizations of the transmitted radio waves and the transmission paths involved in the system. It should be understood that the single-path transmission concept,

TABLE 1. *Diversity measurements*

Run	Test conditions	Transmission paths (See fig. 9)	Polarization		Cross-correlation coefficient of received signal amplitudes				
			Trans.	Rec.	τ, sec				
					0	5	10	−5	−10
5	One transmitting antenna, two receiving antennas; diverging path transmission.	A & C	H	H & H	0.14	0.06	−0.10	0.11	0.03
9	Transmission	A & C	H	H & H	.08	−.05	−.10	.09	.10
25	Two transmitting antennas, one receiving antenna; converging path transmission.	A & B	H & V	HV	.20	.21	.22	.30	−.05
24	One transmitting antenna, one receiving antenna; crossed-polarization one path transmission.	A	HV	HV	[a]1.0				
26	Two transmitting antennas, two receiving antennas; parallel-paths transmission.	A & D	H & V	H & V	0.13	.05	−.03	.06	−.02
28	Two transmitting antennas, two receiving antennas; crossed-path transmission.	B & C	H & V	V & H	.57	−.15	.10	0	−.11

[a] Obtained by visual inspection of paper chart records.

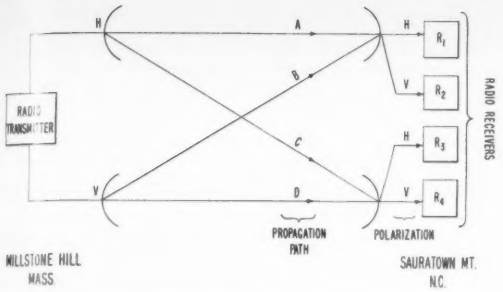


FIGURE 9. *Block diagram of diversity system.*

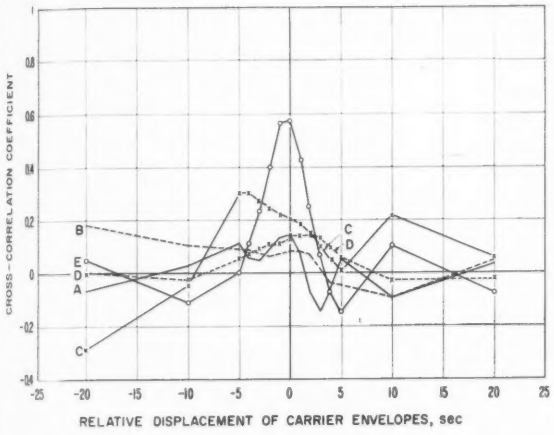


FIGURE 9a. *Cross correlation of tropospheric carrier envelopes versus time displacement of the envelopes relative to each other*

Transmission from Millstone Hill, Mass.
Reception at Sauratown Mt., N.C.
Frequency=417 Mc
Two antennas at each end of system
Antennas, 120-ft square paraboloids
Transverse spacing of antennas, 150 h
July 1958

Run	Curve	Path configuration, from transmitter to receiver
5	A	Diverging paths
9	B	Diverging paths
25	C	Converging paths
26	D	Parallel paths
28	E	Crossed paths

as outlined in figure 9, is an oversimplification of the nature of radio propagation but it is a convenient method of analysis in terms of the geometrical arrangement of the transmitting and receiving antennas.

From table 1 and also figures 9 and 9a it can be seen that diversity is more effective (lowest correlation coefficient) for the cases where separate transmission paths are involved, that is, where no common volume of atmosphere exists at the path midpoint for two or more of the paths. Hence, space diversity would seem to be more effective for the "diverging-path" (or "converging-path") type of system than for the "crossed-path" system since the cross-correlation coefficients were 0.08 to 0.20 and 0.57, respectively. The cross-correlation coefficient for the parallel-path system is 0.13 which compares well with the diverging-path system.

The results shown in figure 9a indicate that radio wave propagation over the crossed paths B and C of figure 9 is not strictly independent and hence the system does not involve four mutually independent paths. However, the cross-correlation coefficient of approximately 0.6 for two of the four paths would still result in radio reception closely approximating that of quadruple diversity [6].

Curve C in figure 9a shows a slight numerical increase in the cross-correlation coefficient for time lags of minus 5 sec and minus 20 sec; these time lags correspond to transverse drift velocities of the common scatter volumes at the path midpoint of approximately 25 mph and 6 mph, respectively; these figures being obtained from the geometry of the paths.

The crossed-polarization single-path test (run 24) gave a cross-correlation coefficient very close to unity. These results indicate that the propagation path fading was independent of the polarization of the radio waves. A close study of the paper chart records for the crossed-polarization tests showed also that the two received carrier envelopes varied identically when the carriers were reflected from an airplane surface.

3.4. Teletype and Binary Signal Errors

As explained previously, the teletype and binary signals at the output of the demodulator (see fig. 1) were recorded as frequency modulation of a 13.02-kc subcarrier on magnetic tape and were also recorded (printed out) by means of a teletype printer at the receiving site. The magnetic tape recorded signals were then played back in the laboratory (at Boulder) into a limiter-discriminator-teletype printer system and the teletype character errors were obtained from the teletype printer records.

The AGC action of the receiving system held the 13.02-kc signal essentially constant at the output of the demodulator while varying the noise level at this point; hence the noise-to-signal level ratio was obtained from the magnetic tape playback by measuring the average noise power in a pass band of approximately 4 cps, placed 130 cps above the 13.02-kc subcarrier.

A low-pass filter having an effective cutoff frequency of 100 cps was inserted in the limiter-discriminator output to the teletype printer in order to limit the noise power spectrum to the bandwidth which was required for the teletype signal.

The binary errors were measured by means of the system shown in block form in figure 10. The binary signal from the FM limiter-discriminator was essentially constant in amplitude with time varying amounts of noise power as shown. The low-pass 13-cps (effective) filter was adjusted to have a rise time somewhat longer than the gating period 0.05T;

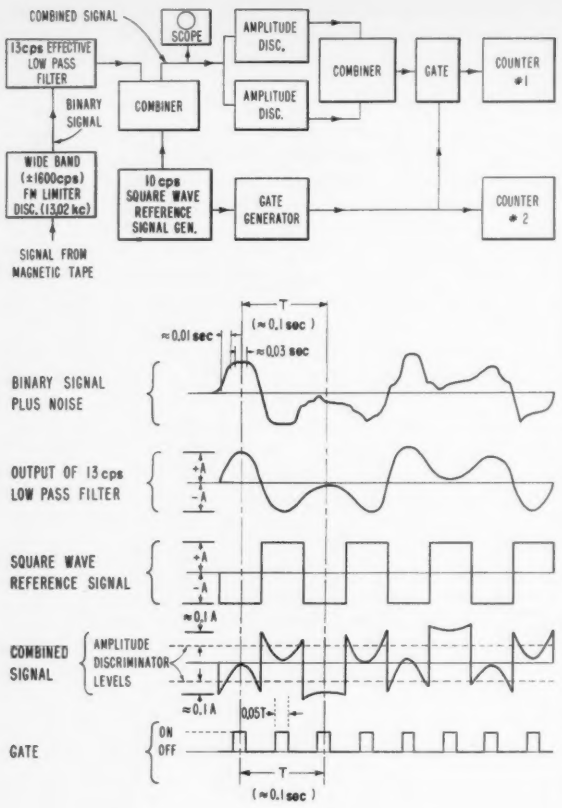


FIGURE 10. *Binary-error measuring system.*

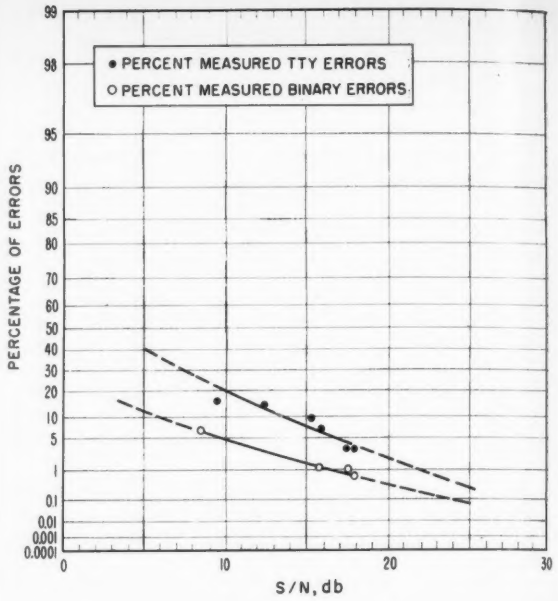


FIGURE 11. *Narrow-band FSK teletype and binary errors versus S/N, for a fading carrier combined with thermal noise.*

S/N is the ratio of the average signal power (FSK 13.02 kc subcarrier) at the limited-discriminator input to the average noise power in a 290 cps effective pass band.

Sauratown Mt., N.C.
July 1958
Transmission from Millstone Hill, Mass.
Signal sampling period 30 min
Carrier frequency = 417.05 Mc
Transmitting and receiving antennas
120-ft square paraboloids

hence the output level of this filter did not change appreciably during the gating period. Triggering levels of the amplitude discriminators were adjusted relative to the combined signal as shown; errors were registered on counter 1 only if the combined signal amplitude exceeded the settings of the discriminator levels during the gating period. In other words, an error was registered if the amplitude of the signal from the 13-cps low-pass filter was less than 0.1A during the gating period. Counter 2 registered the total number of binary pulses occurring during the test period.

From the above it is seen that the errors were measured in terms of whether or not the combined amplitude of the signal plus noise was of the proper polarity and above a particular amplitude ($\approx$0.1A) during the center 10 percent of the signal peak duration. This method of measuring the binary errors conforms to the principles employed by the decision mechanism in the teletype printer.

The results of the teletype and the binary error measurements are shown in figure 11. At this point it should be noted that the effective pass band for the signal plus noise at the input to the limiter-discriminator was 290 cps for both the teletype and the binary signals. It was not feasible to reduce the pass band for the binary signal below 290 cps because of the frequency variation of the subcarrier frequency as obtained from the magnetic tape playback. However, the low-pass filter effective cutoff frequency in the discriminator output circuit for the teletype signal was 100 cps while the effective cutoff frequency of the filter for the binary signal was 13 cps; the cutoff frequencies of these low-pass filters were made only wide enough to pass the binary and the teletype signal spectrums. Hence the ratio of the half passband of the limiter input signal to the low-pass limits of the discriminator output signal was approximately 1.4 for teletype signals and 11 for binary signals. These conditions are close to optimum for the teletype signal but not for the binary signal; that is, the noise-power bandwidth involved in the binary-signal error measurements was much greater than the required binary-signal bandwidth. The net result is that the measured binary errors are increased somewhat in terms of the carrier-to-noise ratio at the limiter input.

Attempts to correlate the measured teletype and binary errors with the short-time nature of the radio-

frequency carrier indicated that the variation of the signal-to-noise ratio as a function of time at the output of the demodulator was dependent somewhat upon the functioning of the demodulator section of the radio receiving system; hence the results of the error measurements are plotted in figure 11 relative to the signal-to-noise ratios at the demodulator output rather than at the IF output. The cumulative distributions of the noise envelope at the demodulator output were measured; these measured distributions were used to determine the average noise power levels at this point. From figure 1 it is evident that when the low-pilot power level is comparable to the noise power level (in the low-pilot channel) the demodulator operates with relatively high noise power at both of its input points and this condition is expected to degrade the signal-to-noise ratio at the output of the demodulator as compared with the carrier-to-noise ratio at the IF output.

4. Conclusions

Measurements of the 417.05-Mc carrier envelope, transmitted over the tropospheric path from Millstone Hill, Mass. to Sauratown Mt., N.C., shows that the cumulative distribution of the instantaneous amplitude of this envelope for 30-min samples roughly approximated a Rayleigh type. Median power-level fade rates of the carrier over this 638-mile path, using one-degree-beamwidth antennas, were less than 0.2 cps. During a small percentage of the time radio signal reflections from airplanes along the path resulted in uniform and large carrier amplitude variations at a rate of up to 7 cps. The measured characteristics of the carrier envelope indicate that the received resultant carrier, over this tropospheric path, resembles a noise signal over a very narrow band of frequencies. It appears as though the received carrier could be closely duplicated in the laboratory by passing a random noise signal through a bandpass filter; this filter would be centered at 417 Mc and its pass characteristics, on each side of the center frequency, would be as shown in figure 7.

Path loss measurements checked within 3.5 db of the calculated values, which were based on the work of P. L. Rice [7] et al.

Dual diversity measurements indicated that polarization diversity was ineffective over this path. It was also found that the diverging-path type of diversity is slightly more effective than the crossed-path type. Angular diversity, involving the use of multiple antenna feeds to produce slightly diverging antenna beams, should be investigated in this system.

Nondiversity error measurements as shown in figure 11 indicated that an average signal power to average noise power ratio (at the limiter-discriminator input) of approximately 20 db gave a 2-percent teletype-character error rate and a corresponding binary error rate of approximately 0.3 percent. These error rates apply to a fading signal and an effective pass band at the limiter input of 290 cps. For optimum bandwidths the required signal-to-noise ratios would be determined by applying bandwidth corrections; the error rate thus obtained from figure 11 would be slightly pessimistic.

The results obtained during these tests give an accurate picture of the nature of the received tropospheric carrier as applied to this particular system and time. Furthermore, it is probable that these results roughly define the limits of power levels, fade rates, spectrum, etc., for other times of the day and the year for this system.

The authors are indebted to M. J. Ogas of the USAF for arranging for these tests, and to Lincoln Laboratory and General Electric Co. personnel for granting permission to carry out these tests. In particular we convey our appreciation to: G. L. Mellen, W. E. Morrow, James H. Chisholm, Donald G. Brennan, J. B. Long, Burt E. Nichols, Don MacLellan, and others of the Lincoln Laboratory; Frank Bias, A. Clark Gunn, Alex Alviros, T. G. Custin, George Rose, and others of the General Electric Co.

We also thank Herbert Wallace and William Collins of the Page Communication Co., and Sergeant John W. Samuelson of the USAF and his group for their assistance in carrying out the tests.

Spectrum analysis work was done by Paul I. Wells and James A. Sykes. Cross correlation analysis of the diversity records was carried out by Harris Janes, A. W. Kirkpatrick, and Mrs. Vivian Fuller. Transmission-loss calculations were made by A. P. Barsis.

5. References

[1] K. A. Norton, P. L. Rice, and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43,** No. 10, 1488 (1955).

[2] A. D. Watt and E. L. Maxwell, Measured statistical characteristics of vlf atmospheric radio noise, Proc. IRE **45,** No. 1, 55 (1957).

[3] A. F. Barghausen, H. B. Janes, F. O. Guiraud, S. Murahata, and C. F. Peterson, Measurements of the power spectra of fading in tropospheric scatter propagation (unpublished report).

[4] G. L. Mellen, W. E. Morrow, A. J. Pote, W. H. Radford, and J. B. Wiesner, Uhf long-range communications systems, M.I.T. Tech. Rept. No. 91, 5 (1955).

[5] T. Hagfors and B. Landmark, Simultaneous variation of amplitude and phase of gaussian noise, with applications to ionospheric forward-scatter signals, Proc. Inst. Elec. Engrs., Pt. B **105,** 555 (1958).

[6] H. Staras, Diversity reception with correlated signals, J. Appl. Phys. **27,** 93 (1956).

[7] P. L. Rice, A. G. Longley and K. A. Norton, Prediction of the cumulative distribution with time of ground wave and tropospheric wave transmission loss (to be published).

Boulder, Colo. (Paper 64D2–46)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March–April 1960

Impedance of a Corner-Reflector Antenna as a Function of the Diameter and Length of the Driven Element

A. C. Wilson*

(September 21, 1959)

Impedance measurements have been made for a monopole in a corner reflector over an image plane as a function of the monopole length, diameter, and position within the corner-reflector structure. The results are presented as a family of curves which should be useful in the design of the driven element for a corner-reflector antenna of the size described in this paper and for other corner-reflector antennas with similar parameters.

1. Introduction

In connection with an NBS-sponsored project to study meteor multipath, a requirement arose for an antenna having a half-power beamwidth in the E-plane of approximately 50° and a capability of radiating a 6-megawatt peak power at a frequency of 40 Mc. On the basis of past experience, a corner-reflector antenna driven by a half-wave dipole was selected for the application.

On the basis of previously measured gains [1][1] and radiation patterns [2] obtainable for corner-reflector antennas having various combinations of lengths and widths, a corner reflector having a length of reflecting planes, $L/\lambda=1.0$, width of reflecting planes, $W/\lambda=1.3$, and an aperture angle of 110° was selected. From figure 48 of reference [2] the beamwidth can be seen to be between 48° and 50° as required. The "first" dipole position, distance from apex of corner reflector, for an aperture angle of 110° can be found in figure 3 of reference [1] to be in the range of 0.2λ to 0.35λ. The gain of a corner-reflector antenna with these parameters should be approximately 12.2 db over an isotropic radiator.

In order to design the driven element for the corner-reflector antenna to radiate the required 6-megawatt peak power, a knowledge of the resistance and reactance values as a function of diameter, length of the driven element, and position within the corner-reflector structure was required.

Theoretical computations of impedance of an antenna with a corner reflector almost always assume reflecting surfaces infinite in extent and aperture angles equal to 45°, 60°, 90°, etc. Assuming perfectly conducting reflecting sheets of infinite extent, the method of images can be applied to obtain an approximation of the corner-reflector antenna impedance [3, 4]. Assuming a dipole position (distance from apex of corner reflector) to be 0.5λ, the distance from the dipole to the reflecting planes may be computed to be 0.205λ for a 110° aperture angle. Using a value of 0.205λ from dipole to reflecting planes for purposes of comparison, the impedance of corner-reflector antennas with 90° and 120° aperture angles may be computed to be approximately $41+j96.5$ and $63+j114$ ohms respectively. For these computations the dipole length was assumed to be one-half wavelength in free space and infinitesimally thin. The effect of the aperture angle can be seen to be significant.

Since a knowledge of the impedance of the corner-reflector antenna as a function of the dipole diameter and length was desired, and also since the parameters of the reflecting surfaces of the corner-reflector antenna selected for this application are small compared with the length of the antenna and an aperture angle of 110° was used, it was necessary to obtain the impedance values experimentally. The results of these measurements are presented in this paper.

2. Experimental Procedure

A scaled experimental model of this antenna was constructed for operation at 400 Mc (a scaling factor of 10). The measurements were made by using a monopole in a corner reflector on an image plane with a corner-reflector width $W/\lambda=0.65$ (one-half 1.3λ). Figure 1 shows a view of the corner reflector on the image plane. The monopoles used in the measurements were made using eight different sizes

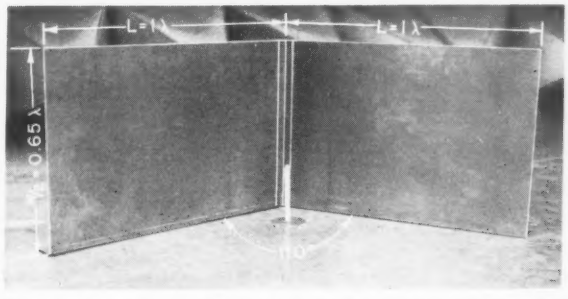


FIGURE 1. *A view of the corner-reflector and monopole on the image plane.*

*Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[1] Figures in brackets indicate the literature references at the end of this paper.

of silver-plated brass tubes fitted with rounded plugs in the free end and a 45° taper plug in the driven end. The diameters of the monopoles used were as follows:

Diameter in inches	Diameter in wavelengths, D/λ	Diameter at 40 Mc in inches	Result curves
1/8	0.00424	1.25	A
9/16	0.00636	1.875	B
1/4	0.00848	2.50	C
3/8	0.0127	3.75	D
1/2	0.0170	5.00	E
3/4	0.0254	7.50	F
1	0.0339	10.00	G
1 1/4	0.0424	12.50	H

The resistance and reactance values for the monopoles at a fixed position in the corner reflector, and with variable length, were determined, using a commercially available impedance-plotting device.

3. Results

Figure 2 presents the resistance and reactance values which were measured for a monopole diameter, $D/\lambda=0.0170$, with monopole lengths l/λ of 0.247, 0.225, 0.212, and 0.200 with the monopole position, distance from apex, variable over the range of 0.15 to 0.4λ. The reactive component was measured to be nearly zero for a monopole length of 0.212λ and a monopole position of 0.25λ. Since these measurements (fig. 2) were obtained for a monopole diameter in the midrange or average of the eight different monopole diameters to be tested, the monopole position for zero reactance should not differ greatly for the other monopole diameters. A monopole position of 0.25λ from apex was used for the remaining measurements of resistance and reactance.

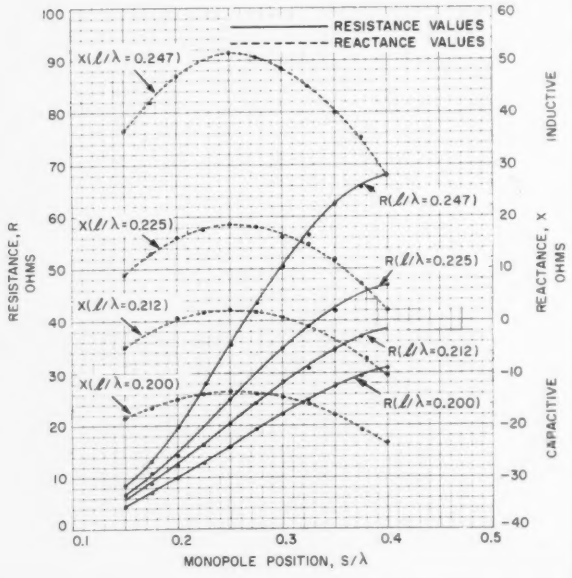


FIGURE 2. *Resistance and reactance curves for a monopole diameter, $D/\lambda=0.0170$, and monopole lengths, $l/\lambda=0.200$ to 0.247, as a function of monopole position, S/λ, in a corner reflector on the image plane.*

Figures 3, 4, and 5 present the curves of resistance and reactance values measured for the monopoles (*A* through *H*) with monopole lengths variable from 0.5λ down to 0.161λ at the fixed monopole position of 0.25λ. (Intermediate values of D/λ can be obtained by interpolating between the curves.)

A corner-reflector antenna has been constructed and is in use for meteor multipath studies. A dipole diameter, $D/\lambda=0.01526$, was chosen for strength availability of standard-size aluminum tubing. On the basis of the measurements described in this report, the remaining corner-reflector antenna parameters were selected as follows: dipole length $l/\lambda=0.42$, and dipole distance from apex=0.25λ.

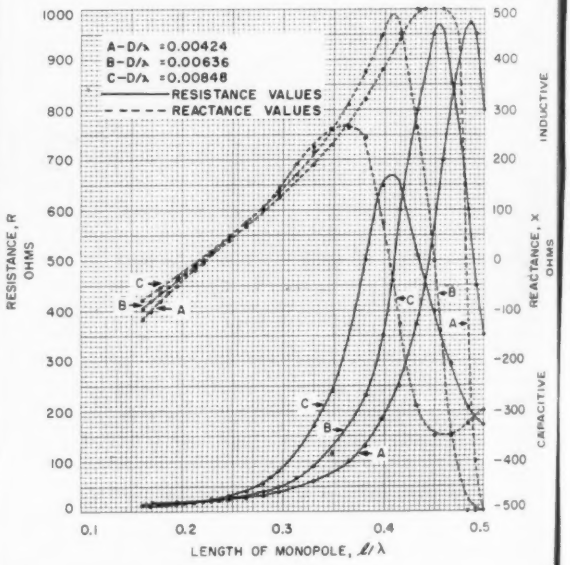


FIGURE 3. *Resistance and reactance curves for different monopole diameters in a corner reflector as a function of length of monopole, l/λ.*

4. Conclusions

This paper is intended to record the basis for the design of the antenna for a meteor multipath measurement project. The measurements presented in this report may also serve as a guide as to what can be expected for corner-reflector antennas with similar parameters. The dipole element diameter may be selected as required for size, strength, power to be radiated, etc., and the dipole length determined.

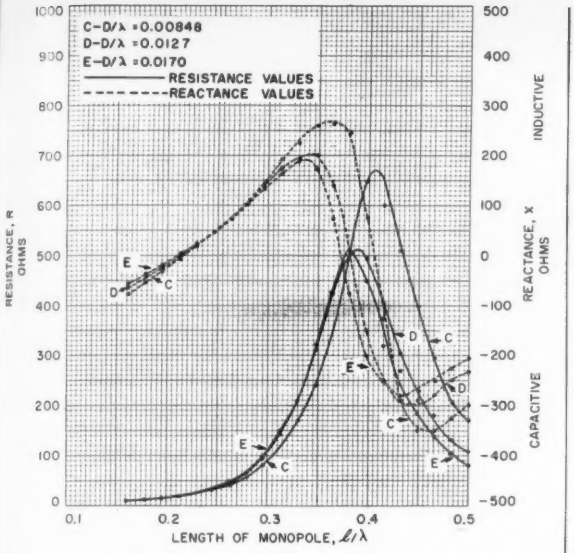


FIGURE 4. *Resistance and reactance curves for different monopole diameters in a corner reflector as a function of length of monopole, l/λ.*

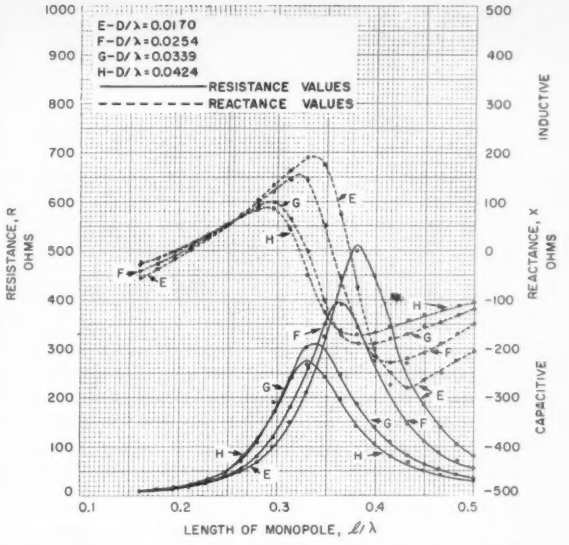


FIGURE 5. *Resistance and reactance curves for different monopole diameters in a corner reflector as a function of length of monopole, l/λ.*

This work was carried out in support of the National Bureau of Standards Meteor Multipath Project at the request of, and under the general guidance of, R. J. Carpenter and G. R. Ochs. The measurements required for this paper were carried out by V. L. Meier.

BOULDER COLO. (Paper 64D2–47)

5. References

[1] H. V. Cottony and A. C. Wilson, Gains of finite-size corner-reflector antennas, IRE Trans. on Ant. Prop. AP–6, 366 (1958).
[2] A. C. Wilson and H. V. Cottony, Radiation patterns of finite-size corner-reflector antennas (to be published).
[3] R. W. P. King, The theory of linear antennas, 370 (Harvard University Press, Cambridge, Mass., 1956).
[4] J. D. Kraus, Antennas, 251, 328 (McGraw-Hill Book Co., Inc., N.Y., 1950).

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March–April 1960

The Electric Field at the Ground Plane Near a Top-Loaded Monopole Antenna with Special Regard to Electrically Small L- and T-Antennas

H. Lottrup Knudsen[1] and T. Larsen

(June 29, 1959; revised September 16, 1959)

The present article deals with the calculation of the electric field strength at the ground plane near electrically small top-loaded antennas having a known current distribution, with special reference to L- and T-antennas. The formulas and numerical values obtained here for this component may be used in calculating the contribution to the ground losses around an antenna of the above-mentioned type due to the vertical component of the earth current.

An exact expression involving an integral has been obtained for the electric field strength at the ground plane due to the current in a linear antenna having an arbitrary inclination. If the length and the height of the antenna is small compared to the wavelength, and if the current distribution on the antenna can be expressed by a finite number of terms of a power series, it is theoretically possibie to obtain a closed expression for the field at the ground plane. However, only in special cases does this expression become sufficiently simple to be of practical value for numerical calculations.

Working formulas have been obtained and numerical calculations carried out for the near zone field of an electrically small vertical or horizontal antenna with a linear current distribution. Based on these results, a calculation has been made of the electric field strength at the ground plane near electrically small L- and T-antennas. Also the relative contribution to this component due to the top loading has been calculated.

1. Introduction

In calculating the ground losses around an antenna, it is customary to assume that the current is flowing in an infinitely thin layer at the surface of the ground so that the current distribution for the purpose of loss calculation may be completely described by a surface current density. However, as has been pointed out by Wait [1],[2] the vertical component of the ground current may in some cases contribute considerably to the ground losses. This is particularly the case if the antenna has a large top loading so that a large current will flow from the top loading to the ground. If, further, the antenna is placed on an island of the same order of magnitude as the top loading so that the ground losses outside the aperture of the antenna are insignificant, the contribution to the ground losses coming from the vertical component of the current distribution may be expected to become still more predominant.

The vertical component of the current density in the earth at the surface of the ground is proportional to the vertical component of the electric field strength at the surface. (With perfect ground conductivity the electric field strength has only a vertical component, and in the case of finite ground conductivity the electric field may often with good approximation be calculated as if the ground was perfectly conducting.) In order to be able to calculate the contribution to the ground losses arising from the vertical component of the current density, we must therefore know the vertical component of the electric field strength at the surface of the earth. The calculation of this component is the object of the present article.

The calculations are carried out under the assumption that the current distribution on the antenna and on the top loading is known. We shall, in particular, consider the case where the top loading consists of a few inclined, or as a special case, horizontal wires. In this case the current distribution on the antenna and on the top loading may be expected to be approximately linear. Working formulas will be developed for and numerical calculations made of the electric

[1] Professor of Electromagnetic Theory, The Technical University of Denmark, Østervoldgade 10 G, Copenhagen K, Denmark.
[2] Figures in brackets indicate the literature references at the end of this paper.

field strength at the ground plane near electrically small L-antennas, T-antennas, and also antennas with four horizontal top-loading wires.

All of the antennas which are considered in this paper are made up of a vertical wire and a top loading consisting of straight wires. We will, therefore, start by developing a formula for the vertical component of the electric field strength at the ground plane, $(E_z)_{z=0}$, due to the current in an inclined wire placed above the ground plane and having a known current distribution. This formula may then be used for finding $(E_z)_{z=0}$ for any one of the antennas considered in this article.

The models used in developing the formulas in this paper are the same as those used in an earlier paper by Knudsen [2] dealing with the surface current density near a top-loaded monopole. However, whereas in the former paper the vector potential method was used in finding the surface current density, it has here been found expedient to use the Hertz dipole method in calculating the electric field strength. The notation used in the report on the surface current density [2] has been used here to as large an extent as seemed practical.

2. List of Principal Symbols

a = length of radial wires in top loading,
A_m = coefficient in power series expansion for current distribution function $g(\sigma)$,
b = height over ground plane of one end of inclined wire,
$B = -(\rho/c)\sin\alpha\cos(\phi-\beta)-(b/c)\cos\alpha$,
c = length of inclined wire,
$C = (\rho/c)^2+(b/c)^2$,
d = reference distance (arbitrary),
$\bar{e}$ = normalized electric field strength at ground plane due to current in inclined wire or top-loaded antenna: $\bar{e} = (kd^2/\zeta I_0)\bar{E}$,
$\bar{e}^h$ = normalized electric field strength at ground plane due to current in one horizontal wire,
$\bar{e}^H$ = normalized electric field strength at ground plane due to current in horizontal members of T-antenna,
$\bar{e}^v$ = normalized electric field strength at ground plane due to current in vertical wire,
$\bar{e}^+$ = normalized electric field strength at ground plane due to current in horizontal members of antenna with four top-loading wires,
$f(s)$ = current distribution function for inclined wire in terms of s,
$g(\sigma)$ = current distribution function for inclined wire in terms of σ,
h = length of vertical member of antenna,
$I(s)$ = current on inclined wire,
I_0 = reference current (arbitrary),
I_b = current in vertical member at ground plane,
$k = \omega\sqrt{\mu\epsilon}$, propagation constant,
n = number of wires in top loading,
$N = C-B^2$,
$\bar{R}$ = vector from point of antenna to point of ground plane $(\rho,\phi,0)$,
$\hat{R}$ = unit vector coparallel to $\bar{R}$: $\hat{R} = \bar{R}/R$,
s = coordinate along inclined wire,
$\hat{t}$ = unit vector pointing in positive direction of inclined wire,
$\hat{x}$ etc. = unit vector in x-direction etc.,
$X = (R/c)^2$,
$X_1 = 1+2B+C$,
α = angle between downwards vertical direction (negative z-axis) and positive direction $(\hat{t})$ of inclined wire,
β = azimuth of inclined wire,
ϵ = dielectric constant,
ζ = intrinsic impedance of free space $\sqrt{\mu/\epsilon}$,
μ = permeability,
$\xi = \rho/h$, normalized distance from base of antenna,
$\rho,\phi,0$ = coordinates of field point at ground plane,
σ = normalized coordinate s/c along inclined wire,
$\tau = a/h$, normalized length of the top-loading wires,
$\psi = \phi-\beta$.

3. Electric Field at Ground Plane Due to Inclined Wire

Let us consider a straight piece of wire of the length c with one of its end points located on the z-axis of a rectangular coordinate system (x,y,z) at a height b above the horizontal (x,y)-plane, the wire forming an angle α with the negative z-axis, and the vertical plane containing the wire forming the angle β with the (z,x)-plane as shown in figure 1. A coordinate s along the wire is introduced, $s=0$ corresponding to the point of the z-axis. We assume that the wire carries a current $I(s)=I_0 f(s)$ where I_0 is an arbitrary reference current. We introduce the normalized coordinate σ defined by $\sigma=s/c$ and the function $g(\sigma)$ defined by $g(\sigma)=g(s/c)=f(s)$.

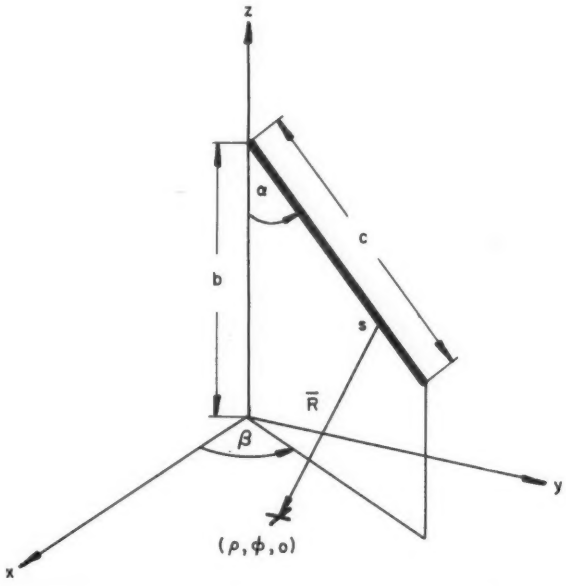


FIGURE 1. *Notation used in calculating the electric field from the current in an inclined wire above a ground plane.*

Let us consider an element ds of the wire located at s. Introducing a unit vector $\hat{t}$ pointing in the positive direction of the wire, this element may be considered a Hertz dipole with the moment $I(s)\,ds\,\hat{t}$. Further, denoting by $\overline{R}$ the vector from a point s on the antenna to a point on the ground plane with the coordinates $(\rho,\phi,0)$, we may express the z-component of the electric field strength $d\overline{E}$ of the Hertz dipole and its image by [3]

$$dE_z=\frac{iI(s)ds e^{ikR}}{2\pi\omega\epsilon}\left[-\frac{k^2}{R}\hat{R}\times(\hat{R}\times\hat{t})-\frac{ik}{R^2}(3\hat{R}\cdot\hat{t}\hat{R}-\hat{t})+\frac{1}{R^3}(3\hat{R}\cdot\hat{t}\hat{R}-\hat{t})\right]_z. \tag{1}$$

In this formula $\hat{R}=\overline{R}/R$ denotes a unit vector coparallel to $\overline{R}$.

Using the reference current I_0 introduced above and further introducing the arbitrary reference distance d; then given an electric field strength $\overline{E}$ we may define a corresponding normalized, dimensionless electric field strength $\overline{e}$ by

$$\overline{e}=\frac{kd^2}{\zeta I_0}\overline{E} \tag{2}$$

where ζ denotes the intrinsic impedance of free space

$$\zeta=\sqrt{\frac{\mu}{\epsilon}}.$$

We thus find the following expression for the vertical component e_z of the normalized electric field strength of the field produced by the wire

$$e_z=\frac{i\left(\frac{d}{c}\right)^2}{2\pi}\int_0\left\{-\frac{(kc)^2}{\left(\frac{R}{c}\right)^3}\left[\frac{R_x}{c}\frac{R_z}{c}t_x-\left(\frac{R_x}{c}\right)^2t_z-\left(\frac{R_y}{c}\right)^2t_z+\frac{R_y}{c}\frac{R_z}{c}t_y\right]\right.$$

$$-\frac{ikc}{\left(\frac{R}{c}\right)^2}\left[\frac{3\frac{R_z}{c}}{\left(\frac{R}{c}\right)^2}\left(\frac{R_x}{c}t_x+\frac{R_y}{c}t_y+\frac{R_z}{c}t_z\right)-t_z\right]$$

$$\left.+\frac{1}{\left(\frac{R}{c}\right)^3}\left[\frac{3\frac{R_z}{c}}{\left(\frac{R}{c}\right)^2}\left(\frac{R_x}{c}t_x+\frac{R_y}{c}t_y+\frac{R_z}{c}t_z\right)-t_z\right]\right\}g(\sigma)e^{ikc\frac{R}{c}}\,d\sigma \quad (3)$$

where the subscripts x,y,z indicate the x-, y-, and z-components of the corresponding vectors.

In the remaining part of this report we will assume that the length and the height of the inclined wire are small compared to the wavelength, i.e., $k(b+c)\ll 1$, and we will only consider the near zone field, i.e., we assume $k\rho\ll 1$.

In this case we get

$$e_z\cong\frac{i\left(\frac{d}{c}\right)^2}{2\pi}\int_0^1\frac{1}{\left(\frac{R}{c}\right)^3}\left[\frac{3\frac{R_z}{c}}{\left(\frac{R}{c}\right)^2}\left(\frac{R_x}{c}t_x+\frac{R_y}{c}t_y+\frac{R_z}{c}t_z\right)-t_z\right]g(\sigma)d\sigma. \quad (4)$$

In order to evaluate the integrals occurring in the above expression it is expedient to expand the known current distribution function $g(\sigma)$ in a power series,

$$g(\sigma)=\sum_{m=0}^{\infty}A_m\sigma^m. \quad (5)$$

Approximating $g(\sigma)$ by a finite number of terms in the above series expression, it is possible to express the integral in the formula above by known functions.

We shall here interest ourselves especially for the case where the top loading consists of only a few wires as in the case of the L-antenna and the T-antenna, respectively. The current distribution may then be assumed to be approximately linear, so that the current distribution function $g(\sigma)$ may be expressed by the two first terms in the above power series. As we are interested mainly in the contribution to the field from the top loading, regarding the current in the vertical member, we shall make the further assumption that it is constant.

4. Vertical Wire

The contribution e_z^v to the vertical component of the electric field at the ground plane due to the vertical member in a top-loaded antenna is considered here. The length of the vertical member is denoted by h and the length of each of the top-loading wires by a. As the reference current I_0 occurring in the above formula, we use here and in the following the current I_b at the bottom of the vertical wire. Approximating the current in the vertical wire by a

constant current, we then find the following expression for the current distribution function $g(\sigma)$

$$g(\sigma)=\frac{I(0)+I(h)}{2I_b}=\frac{1}{2}\left[1+\frac{a}{h+a}\right]=1-\frac{h}{2(h+a)}. \tag{6}$$

In this section and in the following sections we use the length h of the vertical member as the reference distance d in the formulas developed above. We find the following expression for e_z^v

$$e_z^v=\frac{i}{2\pi}\left[1-\frac{h}{2(h+a)}\right]\int_0^1\frac{1}{\left(\frac{R}{h}\right)^3}\left[\frac{3\sigma^2}{\left(\frac{R}{h}\right)^2}-1\right]d\sigma, \tag{7}$$

$$\left(\frac{R}{h}\right)^2=X=\sigma^2+\left(\frac{\rho}{h}\right)^2.$$

In evaluating the above integral we use the expressions for integrals of the type

$$\int_0^1\frac{\sigma^n}{X^m}d\sigma$$

listed in the appendix. We then obtain

$$e_z^v=-\frac{i}{2\pi}\left[1-\frac{1}{2\left(1+\frac{a}{h}\right)}\right]\frac{1}{\left[1+\left(\frac{\rho}{h}\right)^2\right]^{3/2}}. \tag{8}$$

In figure 2, $i\,e_z^v$, i.e., the imaginary unit times the vertical component of the normalized electric field strength at the ground plane, has been plotted as a function of ρ/h and with a/h as a parameter.

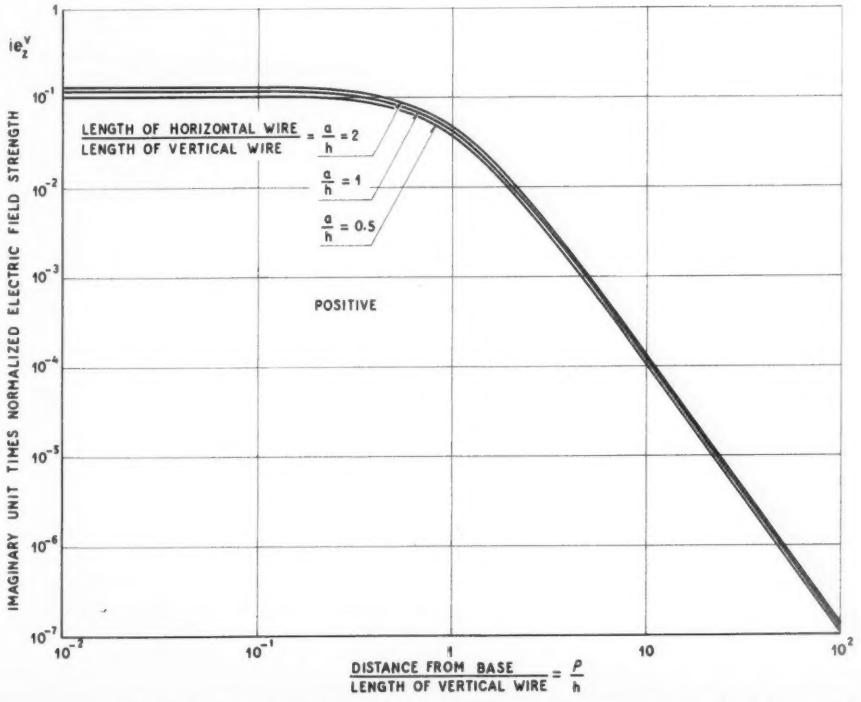


FIGURE 2. *Vertical electric field at ground plane due to vertical current of a top-loaded monopole.*

5. Horizontal Top Loading

We calculate here the contribution e_z^h to the vertical component of the electric field at the ground plane due to one of the top-loading wires of length a in the antenna with a vertical member of length h considered in the last section. As was mentioned before, in applying the formulas developed above, we also here use the current I_b at the bottom of the vertical wire as the reference current I_0 and the length h of the vertical member as the reference distance d. Denoting the number of top-loading wires by $n(n=1,2 \text{ or } 4)$ and making use of the assumption of a linear current distribution, we then have the following expression for the current distribution function $g(\sigma)$

$$g(\sigma)=\frac{1}{n}\frac{I(s)}{I_b}=\frac{1}{n}\frac{a}{a+h}\left(1-\sigma\right). \tag{9}$$

We then find the following expression for e_z^h in the near zone field

$$e_z^h=-\frac{i3}{2\pi n\left(\frac{a}{h}\right)^2\left(1+\frac{a}{h}\right)}\left\{\frac{\rho}{a}\cos(\phi-\beta)\int_0^1\frac{1}{\left(\frac{R}{a}\right)^5}(1-\sigma)\,d\sigma-\int_0^1\frac{\sigma}{\left(\frac{R}{a}\right)^5}(1-\sigma)\,d\sigma\right\}, \tag{10}$$

$$\left(\frac{R}{a}\right)^2=X=\sigma^2-2\sigma\frac{\rho}{a}\cos(\phi-\beta)+\left(\frac{\rho}{a}\right)^2+\left(\frac{h}{a}\right)^2.$$

The integrals occurring in the above formula are evaluated by using the expressions for integrals of the type

$$\int_0^1\frac{\sigma^n}{X^m}\,d\sigma$$

listed in the appendix.

We then find

$$e_z^h=-\frac{i}{2\pi n\left(\frac{a}{h}\right)^2\left(1+\frac{a}{h}\right)N}\left[\frac{1+B}{X_1^{1/2}}-\frac{N+BC}{C^{3/2}}\right]. \tag{11}$$

For large values of ρ it is difficult to calculate e_z^h from this expression with sufficient accuracy as the two terms in the bracket will be nearly equal. In this case it is expedient to expand the two terms in series.

We then get

$$\begin{aligned}ine_z^h\simeq&\frac{\tau^2}{2\pi(1+\tau)(\xi^2\sin^2\psi+1)\xi^2}\left\{\frac{3}{2}\sin^2\psi\cos\psi\right.\\&+\frac{1}{2}\sin^2\psi(-1+5\cos^2\psi)\tau\frac{1}{\xi}+\frac{1}{8}\cos\psi[-18+30\cos^2\psi+\sin^2\psi(-15+35\cos^2\psi)\tau^2]\frac{1}{\xi^2}\\&+\frac{1}{8}[(6-60\cos^2\psi+70\cos^4\psi)\tau+\sin^2\psi(3-42\cos^2\psi+63\cos^4\psi)\tau^3]\frac{1}{\xi^3}\\&+\frac{1}{16}\cos\psi[45-105\cos^2\psi+(75-350\cos^2\psi+315\cos^4\psi)\tau^2\\&\left.+\sin^2\psi(35-210\cos^2\psi+231\cos^4\psi)\tau^4]\frac{1}{\xi^4}\right\},\end{aligned} \tag{12}$$

where

$$\frac{\rho}{h}=\xi \qquad \frac{a}{h}=\tau \qquad \phi-\beta=\psi.$$

Using this expression for ine_z^h for large values of ρ and the exact expression for the smaller values of ρ, we plot in figures 3a–b the quantity ine_z^h, i.e., the imaginary unit times the number of horizontal wires times the vertical component of the normalized electric field strength at the ground plane, as a function of ρ/h and with a/h as a parameter for $\phi-\beta=0°$, and 180°. It is seen as well from the expression for e_z^h as by direct inspection that

$$e_z^h(-(\phi-\beta))=e_z^h(\phi-\beta). \tag{13}$$

6. L-Antenna

Let us consider an L-antenna, the vertical member of which has the length h and the horizontal member the length a as shown in figure 4. The contribution e_z^v to the vertical component of the normalized electric field strength at the ground plane due to the vertical member may be obtained directly from figure 2. As the number of top-loading wires is equal to 1, the contribution e_z^h due to the horizontal member may be obtained directly from figure 3. The vertical component e_z of the normalized electric field strength at the ground plane due to the whole antenna may be expressed as

$$e_z=e_z^v+e_z^h.$$

It is obvious that e_z satisfies the same symmetry relation as e_z^h.

It is of interest to calculate the quotient between the contribution e_z^h of the normalized, electric field strength due to the top loading and the contribution e_z^v due to the vertical member. In figure 5 the ratio e_z^h/e_z^v has been plotted as a function of ρ/h and with a/h as a parameter for $\phi-\beta=0°$. It is seen, as might have been expected, that the relative magnitude of the contribution from the horizontal member decreases with increasing distance from the antenna.

7. T-Antenna

We consider a T-antenna of the height h and with its horizontal member having the length $2a$ as shown in figure 6. The contribution e_z^v to the vertical component of the normalized electric field strength at the ground plane due to the vertical member may be obtained directly from figure 2. Letting β_1 denote the azimuth of one of the two horizontal wires of which the top loading consists, we have for the azimuth of the other radial wire

$$\beta_2=\beta_1+\pi.$$

Utilizing the symmetry properties of the vertical component e_z^h of the normalized electric field strength at the ground plane due to one wire with linearly tapered current distribution, we then find the following expression for the vertical component e_z^H of the normalized electric field strength at the ground plane due to the total top loading by adding the contributions from each of the two wires

$$e_z^H(\phi-\beta_1)=e_z^h(\phi-\beta_1)+e_z^h(\phi-\beta_2)=e_z^h(\phi-\beta_1)+e_z^h(\phi-\beta_1-\pi)=e_z^h(\phi-\beta_1)+e_z^h(\pi-(\phi-\beta_1)). \tag{14}$$

As the number of top-loading wires is equal to 2, we may find $ie_z^h(\phi-\beta_1)$ as ½ times the reading of the curves plotted in figure 3. It is easily verified that $e_z^H(\phi-\beta_1)$ satisfies the following symmetry relations

$$e_z^H(-(\phi-\beta_1))=e_z^H(\phi-\beta_1) \tag{15}$$

$$e_z^H(\pi-(\phi-\beta_1))=e_z^H(\phi-\beta_1). \tag{16}$$

The vertical component e_z of the normalized electric field strength at the ground-plane due to the whole antenna may now be obtained from

$$e_z=e_z^v+e_z^H.$$

It is obvious that e_z satisfies the same symmetry relations as e_z^H.

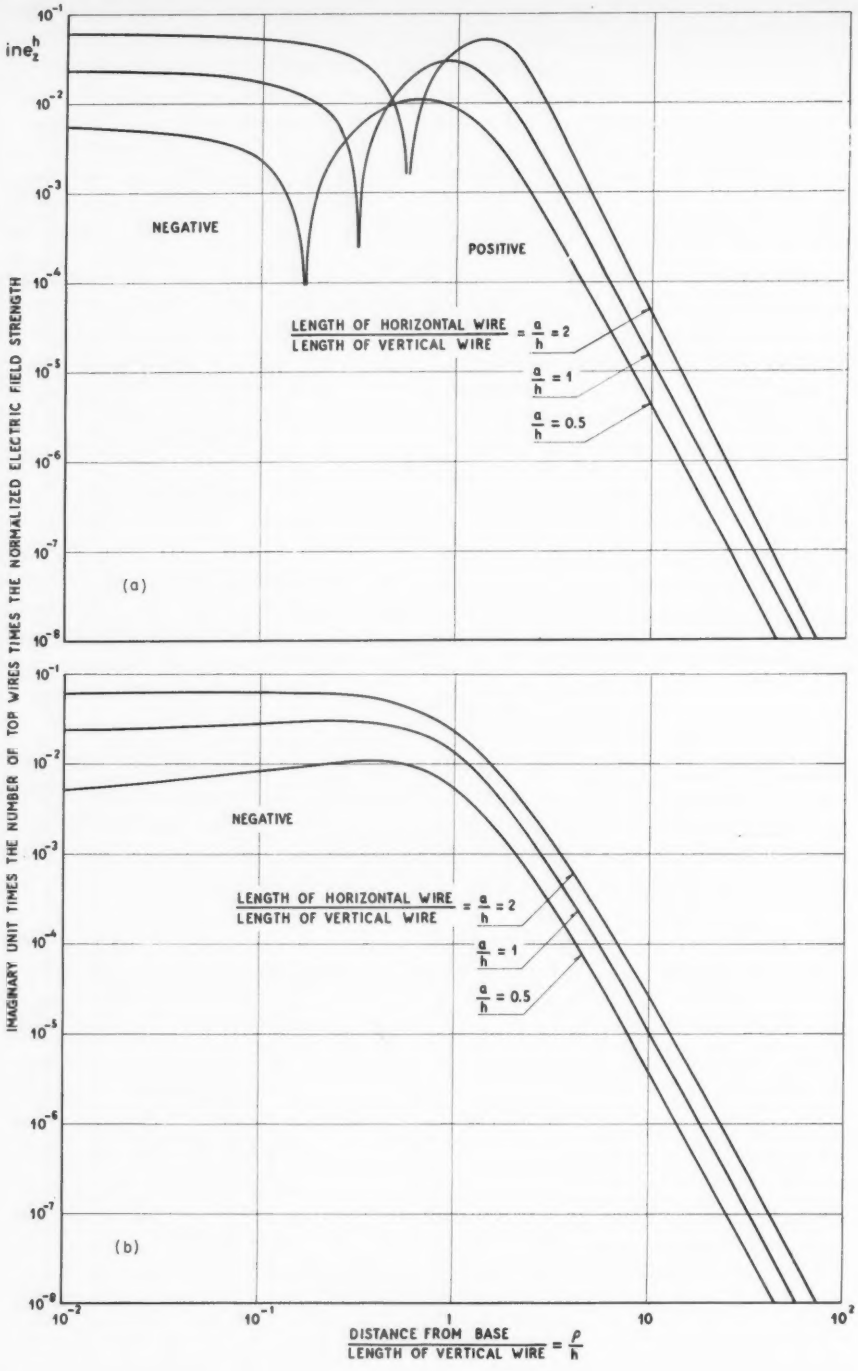

IMAGINARY UNIT TIMES THE NUMBER OF TOP WIRES TIMES THE NORMALIZED ELECTRIC FIELD STRENGTH
ine_z^h
NEGATIVE
POSITIVE
LENGTH OF HORIZONTAL WIRE / LENGTH OF VERTICAL WIRE = a/h = 2
a/h = 1
a/h = 0.5
(a)
NEGATIVE
LENGTH OF HORIZONTAL WIRE / LENGTH OF VERTICAL WIRE = a/h = 2
a/h = 1
a/h = 0.5
(b)
DISTANCE FROM BASE / LENGTH OF VERTICAL WIRE = ρ/h

FIGURE 3. *Vertical electric field at ground plane due to horizontal current in one top wire of a top-loaded monopole.*

a: $\phi-\beta=0°$, b: $\phi-\beta=180°$.

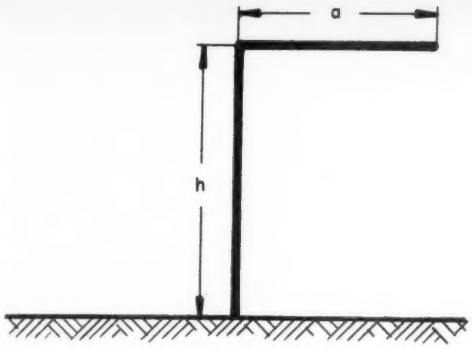


FIGURE 4. L-*antenna*

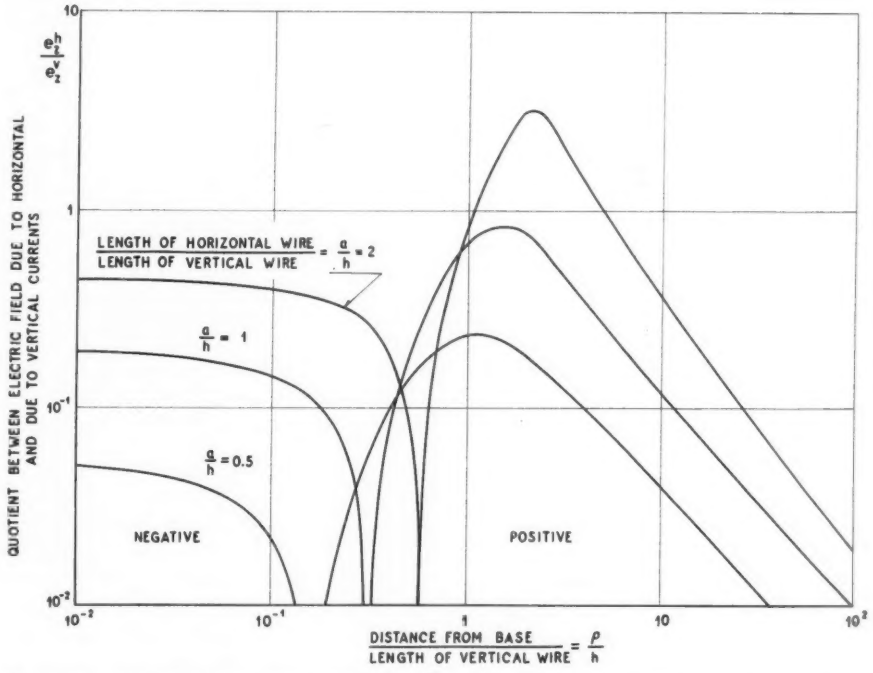


FIGURE 5. *Quotient between the contributions to the vertical electric field at ground plane due to the top loading and due to the vertical member of an* L-*antenna,* $\phi-\beta=0°$.

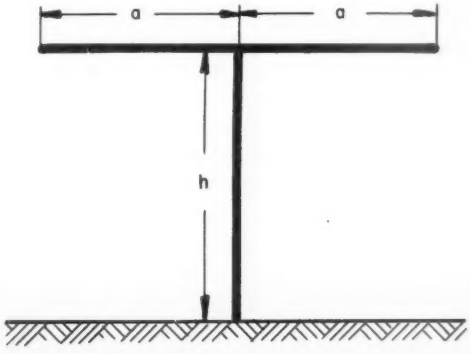


FIGURE 6. T-*antenna.*

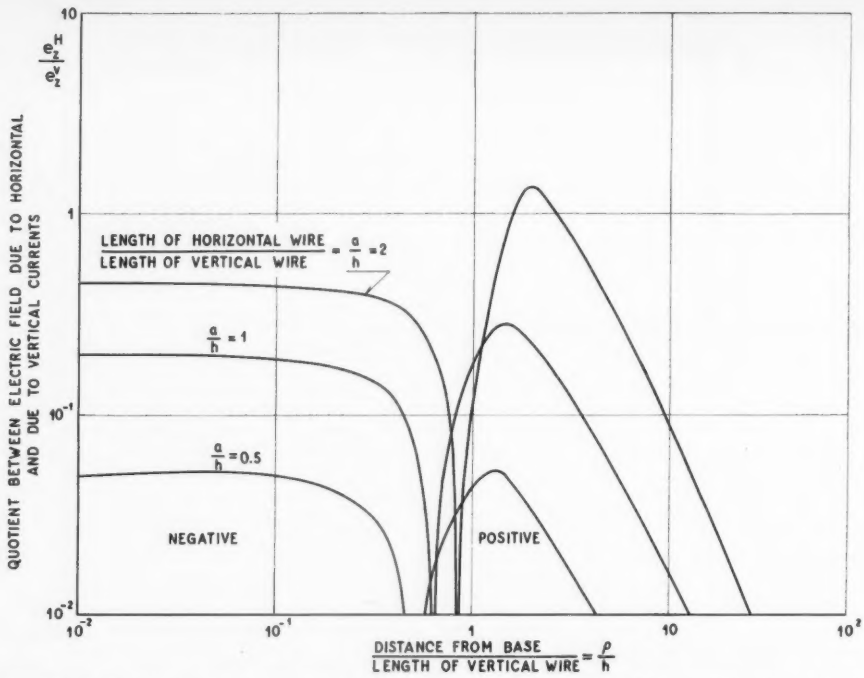


FIGURE 7. *Quotient between the contributions to the vertical electric field at ground plane due to the top loading and due to the vertical member of a* T*-antenna,* $\phi-\beta_1=0°$.

In figure 7 the ratio e_z^H/e_z^v has been plotted as a function of ρ/h and with a/h as a parameter for $\phi-\beta_1=0°$. We see that also in the case of the T-antenna the relative magnitude of the contribution from the top loading decreases with increasing distance from the antenna.

8. Antenna With Four Top-Loading Wires

We consider here an antenna consisting of a vertical wire and four equiangularly spaced, horizontal wires as shown in figure 8. We include this case in the present investigation because it approximates the case of a disk-loaded antenna, so that we are hereby enabled to compare fairly well the formulas and numerical results derived in this paper with the corresponding formulas and numerical results that might be derived later on the basis of the model of a disk-loaded antenna. We assume also in the present case the current distribution on the antenna including the top loading is linear. Whereas this current distribution probably fairly well approximates the actual current distribution in the case of L- and T-antennas, it is questionable whether it does so in the case of the antenna with four top-loading wires. However, we use the assumption of a linear current distribution for the sake of simplicity, remembering that the main emphasis in this paper is on L- and T-antennas.

The length of the vertical wire is denoted by h and the length of each of the radial wires by a. The contribution e_z^v to the vertical component e_z of the electric field strength at the ground plane may be obtained directly from figure 2. Letting β_1, β_2, β_3, and β_4 denote the azimuth of the four wires of which the top loading consists, we have

$$\beta_2=\beta_1+\frac{\pi}{2}$$

$$\beta_3=\beta_1+\pi$$

$$\beta_4=\beta_1+\frac{3\pi}{2}.$$

FIGURE 8. *Antenna with four top-loading wires.*

Using this notation and the symmetry properties of the vertical component e_z^h of the normalized electric field strength at the ground plane due to one wire, we then find the following expression for the vertical component e_z^+ of the normalized electric field strength at the ground plane due to the total top loading.

$$e_z^+(\phi-\beta_1)=e_z^h(\phi-\beta_1)+e_z^h\left(\phi-\beta_1-\frac{\pi}{2}\right)+e_z^h(\phi-\beta_1-\pi)+e_z^h\left(\phi-\beta_1-\frac{3\pi}{2}\right)$$

$$=e_z^h(\phi-\beta_1)+e_z^h\left(\frac{\pi}{2}-(\phi-\beta_1)\right)+e_z^h(\pi-(\phi-\beta_1))+e_z^h\left(\phi-\beta_1+\frac{\pi}{2}\right). \quad (17)$$

As the number of top-loading wires is equal to 4, $ie_z^h(\phi-\beta_1)$ may be found as 1/4 times the value obtained from the curves plotted in figure 3. We may easily verify that $e_z^+(\phi-\beta_1)$ satisfies the following symmetry relations

$$e_z^+(-(\phi-\beta_1))=e_z^+(\phi-\beta_1), \quad (18)$$

$$e_z^+(\pi-(\phi-\beta_1))=e_z^+(\phi-\beta_1), \quad (19)$$

$$e_z^+\left(\frac{\pi}{2}-(\phi-\beta_1)\right)=e_z^+(\phi-\beta_1). \quad (20)$$

The vertical component e_z of the normalized electric field strength at the ground plane due to the whole antenna is then obtained from

$$e_z=e_z^v+e_z^+.$$

The component e_z is seen to satisfy the same symmetry relations as e_z^+.

In figure 9 the ratio e_z^+/e_z^v has been plotted as a function of ρ/h and with a/h as a parameter for $\phi-\beta=0°$.

9. Conclusion

Formulas have been derived and numerical calculations carried out for the vertical component of the electric field strength at the ground plane in the near zone field of a vertical or horizontal linear antenna having a length and height above the ground plane that is small

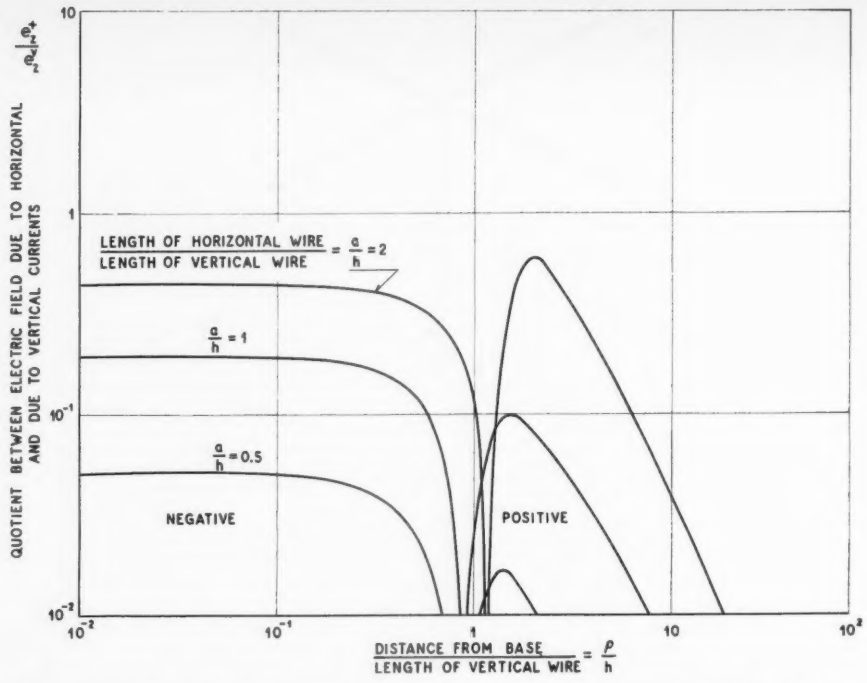


FIGURE 9. *Quotient between the contributions to the vertical electric field at ground plane due to the top loading and due to the vertical member of an antenna with four top-loading wires,* $\phi-\beta_1=0°$.

compared to the wavelength. These results have been used in obtaining the above-mentioned component and the relative magnitude of the contribution to this component due to the horizontal top loading in the near zone field of electrically small L-antennas, T-antennas, and antennas with four horizontal, top-loading wires. The results obtained here should be useful in calculating the contribution to the ground losses due to the vertical component of the earth currents.

It is seen from the curves that there is a reduction of the relative vertical electric field when symmetrical top loading is used. This seems to be an advantage of the T- and four top hot wire antenna over the L-antenna.

Part of this investigation was carried out when H. Lottrup Knudsen was temporarily employed as a consultant at the National Bureau of Standards in Boulder, Colo., in the summer of 1957. Tove Larsen's participation in the investigation was made possible by grants from the Technical University of Denmark and from Den polytekniske Læreanstalts ingeniørvidenskabelige Fond and G. A. Hagemanns Mindefond.

Our thanks are due to James R. Wait for having suggested the problem and for helpful comments, to Erik B. Hansen for having checked the formulas with the result that one formula was considerably condensed, and to Anabeth C. Murphy for having made the numerical computations.

10. Appendix

Some Integrals of the Type $\int \frac{\sigma^n}{(\sigma^2+2B\sigma+C)^m} d\sigma$

In calculating the vertical component of the electric field strength at the ground plane in the near zone field of a vertical or horizontal linear antenna having a linear current distribution, the length and height of which is small compared to the wavelength, we have use for closed expressions for some integrals of the type

$$\int_0^1 \frac{\sigma^n}{X^m} d\sigma$$

where

$$X=\sigma^2+2B\sigma+C,$$

B and C being constants satisfying the inequality

$$N=C-B^2\geq 0.$$

The corresponding indefinite integrals have partly been obtained from Gröbner and Hofreiter's table of integrals [4].

Introducing the following abbreviation

$$X_1=(X)_{\sigma=1}=1+2B+C$$

we have for the pertinent definite integrals

$$\int_0^1 \frac{1}{X^{3/2}} d\sigma=\frac{1}{N}\left[\frac{1+B}{X_1^{1/2}}-\frac{B}{C}\right], \tag{21}$$

$$\int_0^1 \frac{1}{X^{5/2}} d\sigma=\frac{1}{3N}\left[\frac{1+B}{X_1^{3/2}}-\frac{B}{C^{3/2}}\right]+\frac{2}{3N^2}\left[\frac{1+B}{X_1^{1/2}}-\frac{B}{C^{1/2}}\right], \tag{22}$$

$$\int_0^1 \frac{\sigma}{X^{5/2}} d\sigma=\frac{1}{3N}\left[-\frac{B+C}{X_1^{3/2}}+\frac{1}{C^{1/2}}\right]+\frac{2B}{3N^2}\left[-\frac{1+B}{X_1^{1/2}}+\frac{B}{C^{1/2}}\right], \tag{23}$$

$$\int_0^1 \frac{\sigma^2}{X^{5/2}} d\sigma=\frac{1}{3N}\left[\frac{2B^2-C+BC}{X_1^{3/2}}-\frac{B}{C^{1/2}}\right]+\frac{B^2+C}{3N^2}\left[\frac{1+B}{X_1^{1/2}}-\frac{B}{C^{1/2}}\right]. \tag{24}$$

11. References

[1] J. R. Wait, On the calculations of transverse current loss in buried wire ground systems, Appl. Sci. Research, Sec. B **7,** No. 1, p. 81 (1958).

[2] H. Lottrup Knudsen, Earth currents near a top-loaded monopole antenna with special regard to electrically small L- and T- antennas, J. Research NBS **62,** 283 (1959) RP2961.

[3] R. W. P. King, Electromagnetic engineering, Vol. 1 (McGraw-Hill Book Co., Inc., New York, N.Y., 1945).

[4] W. Gröbner and N. Hofreiter, Integraltafel; erster Teil. Unbestimmte Integrale, Springer-Verlag, Wien and Innsbruck (1949).

Boulder, Colo.

(Paper 64D2–48)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March–April 1960

Terrestrial Propagation of Very-Low-Frequency Radio Waves

A Theoretical Investigation

James R. Wait*

(September 11, 1959)

A self-contained treatment of the waveguide mode theory of the propagation of very-low-frequency radio waves is presented. The model of a flat earth with a sharply bounded and homogeneous ionosphere is treated for both vertical and horizontal dipole excitation. The properties of the modes are discussed in considerable detail.

The influence of earth curvature is also considered by reformulating the problem using spherical wave functions of complex order. The modes in such a curved guide are investigated and despite the initial complexity of the general solution, many interesting and limiting cases may be treated in simple fashion to yield useful and convenient formulas for calculation.

Other factors considered are the influence of the earth's magnetic field, antipodal effects, resonator type oscillations, and the influence of stratification at the lower edge of the ionosphere.

1. Introduction

The concept that radio waves are channeled between the earth and the ionosphere as in a waveguide has proved to be very useful at very low frequencies (<30 kc). In 1919, G. N. Watson [1] [1] employed this approach when he considered, at least in a formal way, the propagation of electromagnetic waves between an idealized homogeneous spherical earth and a concentric reflecting layer. Because of the extremely poor convergence of the exact series solutions, Watson devised a technique to convert this to a more rapidly converging series using function-theoretic means. The new representation corresponds to the sum of residues at poles in the complex plane and hence the name "residue series." The waves associated with these poles are the waveguide modes. Watson studied the numerical properties of these modes for the case of long waves or low frequencies and on assumption of a very highly conducting shell. This particular aspect of his investigation was prompted by the recent discoveries of Marconi that radio waves decay much more slowly with distance than predicted on the basis of classical diffraction theory in the absence of a reflecting shell. Watson found that the modes of low attenuation behaved like

$$\frac{1}{(\sin d/a)^{\frac{1}{2}}}\, e^{-\alpha\sqrt{f/\sigma}\, d}$$

where d is the great circle distance, f is the frequency, σ is the conductivity of the reflecting ionosphere, a is the radius of the earth, and α is a constant. For frequencies in the range from about 20 to 40 kc, observed field strengths behaved more or less in this fashion if the effective ionospheric conductivity was taken to be about 10^{-4} mhos/m or a conductivity of the same order as "tap water". Actually for frequencies in this range some 10 to 30 modes would be excited and if the complete mode sum were considered, the calculated field strength versus distance curve using such a model would show many rapid and violent undulations. Such a behavior is not observed under normal conditions and this fact alone is sufficient cause to reject this model even from a phenomenological viewpoint. The same model with certain refinements has been discussed more recently by Rydbeck [2] in a monograph, Bremmer [3] in his book, Schumann [4 to 6] in a series of papers, and most recently by Kaden [7]. From the frequency

* Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[1] Figures in brackets indicate the literature references at the end of this paper.

analysis of atmospheric wave forms [8, 9], it is known that the attenuation rate does not vary like $\sqrt{f}$ except possibly at frequencies near 1 kc. Actually, the attenuation rate decreases with increasing frequency in the range from about 2 to 18 kc and thereafter increases. A behavior of this kind is highly suggestive of a Brewster angle effect. Such a proposal was first made by Namba [10] as far as this writer can ascertain. It thus appears than the ionosphere at vlf is behaving more like a magnetic wall (tangential H near zero) rather than an electric wall (tangential E near zero) as postulated by Watson, Bremmer, and Schumann.

Contributions to the waveguide mode concept have also been made by Budden [11 to 13 incl.] who unlike the early workers did not assume a highly conducting reflective layer at the outset. His model was a vertical electric dipole source in the space between the surface of a flat perfectly conducting ground and a sharply bounded homogeneous ionosphere. Various extensions and generalizations have been made by Al'pert [14, 15], Lieberman [16], Wait [17 to 22 incl.], Howe [23], and Friedman [24]. Their work is referred to from time to time in the text. It is the purpose of the present paper to present a unified treatment of the mode theory of vlf propagation. The results include much of the above work as special cases. In fact, in some instances, the analysis follows the work of Budden, Bremmer, and Rydbeck rather closely, although many new results are derived. While no attempt is made to present numerical results, limiting cases and simplified forms of the general solutions are discussed in some detail. An extensive bibliography is included in this paper for the convenience of those who are more interested in related numerical results and experimental data. It is intended that this paper will serve as a theoretical basis for subsequent papers by A. G. Jean, W. L. Taylor, A. D. Watt, and the author.

2. Basic Concepts

To introduce the subject a very simple model is chosen. The earth and the ionosphere are represented by perfectly conducting planes. In terms of a cylindrical coordinate system (ρ, ϕ, z) the ground surface is the plane $z=0$ and the lower boundary of the ionosphere is the plane $z=h$. The source is now considered to be a vertical electric dipole located on the ground. The electric field observed at some other point on the ground plane has only a vertical component and can be deduced by considering the images of the source dipole. These images are located at $z=\pm 2h$, $\pm 4h$, $\pm 6h$, etc., and all have equal sign and magnitude, because of the assumed perfect conductivity of the bounding walls. These images will always direct a wave broadside since the radiation from each image is in phase. At a distance which is large compared to h, this field can be calculated by replacing the line of dipole images by a continuous line source carrying an equivalent uniform current I_a. This current is the average current along the z axis and is given by

$$I_a=\frac{ds}{h}\,I$$

in terms of the height ds of the dipole and its current I. Now the field of a line source of current is well known and thus

$$E_z=\frac{I_a\mu\omega}{4}\,H_0^{(2)}(k\rho)=\frac{\mu\omega Ids}{4h}\,H_0^{(2)}(k\rho) \tag{2.1}$$

where $H_0^{(2)}(k\rho)$ is the Hankel function of the second kind of argument $k\rho$, μ is the permeability of the space, ω is the angular frequency, and $k=2\pi/\lambda$. When $\rho \gg \lambda$, the Hankel function can be replaced by the first term of its asymptotic expansion and this leads readily to

$$E_z\cong\frac{\eta}{2}\,\frac{Ids}{h(\lambda\rho)^{\frac{1}{2}}}\,e^{i\pi/4}e^{-ik\rho} \tag{2.2}$$

where $\eta=(\mu/\epsilon)^{\frac{1}{2}}\cong 120\pi$. As mentioned above, this field corresponds to the radiation directed broadside so the rays are parallel to the bounding walls. However, there will be other angles where the rays emanating from each of the dipoles in the line of images are also in phase.

Such a resonance condition exists when

$$2hC=n\lambda \tag{2.3}$$

where C is the cosine of the angle subtended by the rays and the z axis and n is an integer (see fig. 1). It is seen that for each value of n there are two families of rays which have the same radial phase velocity (i.e.,$=c/S$) but with opposing vertical phase velocities (i.e.,$=\pm c/C$). Again the radiation of these sets of waves (i.e., modes) can be imagined to originate from an equivalent line source. The strength of this line source is IS where S is the sine of the angle subtended by the rays and the vertical direction. To obtain the resultant vertical field, this must be again multiplied by S. Consequently the resultant field of all the families of rays or modes is obtained by summing over integral values of n from 0 to ∞ to give

$$E_z=\frac{\mu\omega Ids}{4h}\sum_{n=0}^{\infty}\epsilon_n S_n^2 H_0^{(2)}(kS_n\rho) \tag{2.4}$$

where $\epsilon_0=1$, $\epsilon_n=2(n=1, 2, 3\ \ldots)$ and $S_n=(1-C_n^2)^{\frac{1}{2}}$ and $C_n=n\lambda/2h$. The term $n=0$, corresponding to mode zero discussed above, is only included once in the summation, whereas the higher modes are included twice. In the far field, this expression for the field reads

$$E_z\cong\frac{\eta Ids}{2h(\lambda\rho)^{\frac{1}{2}}}e^{i\pi/4}\sum_{n=0,1,2,\ldots}^{\infty}\epsilon_n S_n^{3/2}e^{-ikS_n\rho}. \tag{2.5}$$

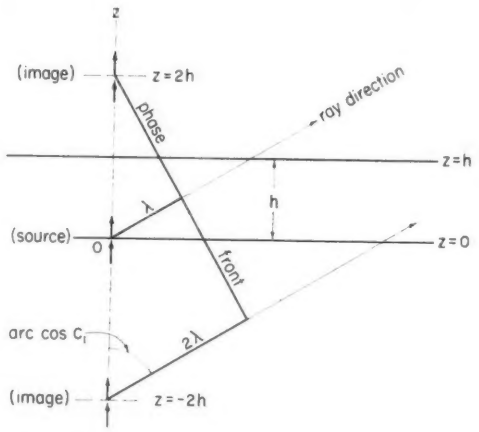


FIGURE 1. *Depicting ray-geometry corresponding to the first mode between parallel plates; for resonance,* $\lambda=2hC_1$.

Up to this point the bounding walls have been assumed to be of perfect conductivity: The reflection coefficients for the rays are always $+1$. Another simple case is when the upper boundary has a reflection coefficient of -1 corresponding to a perfect magnetic conductor, and the lower boundary still has a reflection coefficient of $+1$ corresponding to the more common perfect electrical conductor. For this situation the images are now located at $z=\pm h$, $\pm 2h$, . . . but now they alternate in sign. It may be observed that there is no coherent family of rays directed broadside. This would have been the zero-order mode. The resonance condition for the modes is now

$$2hC=(n-\tfrac{1}{2})\lambda \tag{2.6}$$

where $n=1, 2, 3, \ldots$. The corresponding expression for the vertical electric field is thus given by

$$E_z=\frac{\mu\omega Ids}{2h}\sum_{n=1,2,3,\ldots}^{\infty}S_n^2\,H_0^{(2)}(kS_n\rho) \tag{2.7}$$

where the summation starts at $n=1$ and includes all positive integers. In the far field

$$E_z \cong \frac{\eta I ds}{h(\lambda\rho)^{\frac{1}{2}}} e^{i\pi/4} \sum_{n=1,2,3,\ldots}^{\infty} S_n^{3/2} e^{-ikS_n\rho}. \tag{2.8}$$

In the foregoing discussion the observer and the source, which is a vertical electric dipole, are located on the ground plane. The above results are easily generalized to a finite source height, z_0, and a finite observer height, z, by inserting the factor $\cos(kz_0C_n)\cos(kzC_n)$ inside the summations of eqs (4), (5), (7), and (8). This can be verified by returning to the image picture and noting now that they are located at $z=-h$, $\pm(2h+z_0)$, $\pm(2h-z_0)$, $\pm(4h+z_0)$, $\pm(4h-z_0)$, ... It may also be observed that the $\cos(kzC_n)$ when replaced by $[\exp(ikzC_n)+\exp(-ikzC_n)]/2$ can be identified as a family of upgoing and downcoming rays within the guide.

The important modifications of the preceding formulas as a result of imperfect reflection can be obtained by a rather simple physical argument. The complete treatment requires a more mathematical approach which is to be described later on.

The reflection coefficient for a ray incident on the ground plane at an angle (whose cosine is C) is denoted $R_g(C)$. The corresponding reflection coefficient for the upper boundary which is the lower edge of ionosphere is denoted $R_i(C)$. The resonance condition now has the form

$$R_g(C)R_i(C)e^{-i2khC} = e^{-i2\pi n} \tag{2.9}$$

which reduces to eq (3) if the reflection coefficients are both $+1$ and reduces to eq (6) if one reflection coefficient is $+1$ and the other -1. Physically, the above more general form can be the condition for a ray to traverse the guide twice, be reflected at each boundary, and yet still suffer a net phase shift of $2\pi n$ radians where n is an integer. Since $R_g(C)$ and $R_i(C)$ may be complex and less than unity the value of C (i.e., C_n) which satisfies the resonance equation may also be complex. The angle of incidence of these rays in the guide are thus also complex. The corresponding value of S_n is also complex and this results in attenuation of the wave in the radial direction. In fact, the attenuation constant is minus the imaginary part of kS_n in nepers per unit distance.

When the angle or its cosine C must be complex in order to satisfy a resonance equation, the resulting waves are damped. The numerical solution of such a complex resonance equation is quite difficult, in general, since it is not usually possible to obtain an explicit expression for C in terms of known parameters. This aspect of the problem is discussed in a later section.

3. Formulation for Flat Earth Case

3.1. Vertical Dipole Excitation

It is now desirable to formulate the problem in a more definite fashion. A vertical electric dipole of moment $I\,ds$ is placed in a homogeneous plane layer bounded by two plane interfaces (see fig. 2). The lower interface is at $z=0$ corresponding to the surface of a homogeneous ground of conductivity σ_g and dielectric constant ϵ_g. The upper interface at $z=h$ is the lower edge of a homogeneous ionosphere which for the moment is assumed to be isotropic and has effective electrical constants σ_i and ϵ_i. The fields in these regions can be derived from a Hertz vector which has only a z component, Π_z. Thus, for $h \geqq z \geqq 0$

$$\begin{aligned} E_\rho &= \frac{\partial^2}{\partial\rho\partial z}\Pi_z & H_\rho &= 0 \\ E_\phi &= 0 & H_\phi &= -i\epsilon\omega\frac{\partial\Pi_z}{\partial\rho} \\ E_z &= \left(k^2+\frac{\partial^2}{\partial z^2}\right)\Pi_z & H_z &= 0 \end{aligned} \tag{3.1}$$

where $k=(\epsilon\mu)^{\frac{1}{2}}\omega=2\pi/\lambda$.

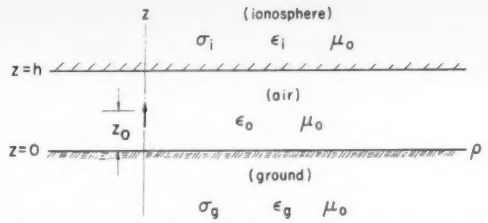


FIGURE 2. *Cylindrical coordinate system for the vertical dipole between the two plane interfaces.*

Similar expressions are applicable for the regions $z>h$ and $z<0$ where the z directed Hertz vectors are denoted $\Pi_z^{(i)}$ and $\Pi_z^{(g)}$, respectively, and the corresponding wave numbers are k_i and k_g, respectively. The formal solution of this problem is obtained in a straightforward fashion by requiring that the tangential field components E_ρ and H_ϕ are continuous across the two plane interfaces. An equivalent statement of these matching conditions is as follows:

$$\left.\begin{aligned} k^2\Pi_z &= k_g^2\Pi_z^{(g)} \\ \frac{\partial\Pi_z}{\partial z} &= \frac{\partial\Pi_z^{(g)}}{\partial z} \end{aligned}\right] \quad \text{for } z=0, \tag{3.2}$$

$$\left.\begin{aligned} k^2\Pi_z &= k_i^2\Pi_z^{(i)} \\ \frac{\partial\Pi_z}{\partial z} &= \frac{\partial\Pi_z^{(i)}}{\partial z} \end{aligned}\right] \quad \text{for } z=h. \tag{3.3}$$

To facilitate the solution, the primary excitation resulting from the source dipole is represented as a spectrum of plane waves. This well-known representation, for the primary Hertz function, is given as follows: [25]

$$\Pi_z^{(p)} = \frac{M\exp\{-ik[\rho^2+(z-z_0)^2]^{\frac{1}{2}}\}}{[\rho^2+(z-z_0)^2]^{\frac{1}{2}}} = \frac{Mik}{2}\int_\Gamma H_0^{(2)}(kS\rho)\exp[-ikC|z-z_0|]\,dC \tag{3.4}$$

where $M=I\,ds/(4\pi i\omega\epsilon)$ and $S=(1-C^2)^{\frac{1}{2}}$. The integration variable C can be regarded as the cosine of the angle of incidence of the plane waves in the spectrum. Γ is the contour of integration and it extends from $-\infty$ along the negative real axis to the origin, then out along the real axis to $+\infty$. It should be noted since C can be greater than unity complex angles in the spectrum occur. The above form for the primary excitation then suggests that the resultant Hertz function for the three regions can be written in the respective forms

(I) $$\Pi_z = \Pi_z^{(p)} + \int_\Gamma [A(C)e^{-ikCz} + B(C)\,e^{+ikCz}]\,H_0^{(2)}\,kS\rho)\,dC \tag{3.5}$$

for $0\leqq z\leqq h$;

(II) $$\Pi_z^{(g)} = \int_\Gamma G(C)e^{+ik_gC_gz}H_0^{(2)}(kS\rho)dC \tag{3.6}$$

for $z\leqq 0$; and

(III) $$\Pi_z^{(i)} = \int_\Gamma I(C)e^{-ik_iC_iz}H_0^{(2)}(kS\rho)dC \tag{3.7}$$

for $z>h$. In the above, it can be verified that these Hertz functions satisfy the appropriate wave equation subject to the conditions that

$$N_g(1-C_g^2)^{\frac{1}{2}} = (1-C^2)^{\frac{1}{2}} = N_i(1-C_i^2)^{\frac{1}{2}} \tag{3.8}$$

where

$$N_g=\left(\frac{\sigma_g+i\epsilon_g\omega}{i\epsilon\omega}\right)^{\frac{1}{2}} \text{ and } N_i=\left(\frac{\sigma_i+i\epsilon_i\omega}{i\epsilon\omega}\right)^{\frac{1}{2}}.$$

Terms containing exp $(-ik_gC_gz)$ and exp (ik_iC_iz) are not permitted since they would violate the Sommerfeld radiation condition at $|z|\to\infty$.

The form of the unknown function $A(C)$, $B(C)$, $G(C)$, and $I(C)$ can be obtained explicitly by using the four equations of continuity. This purely algebraic process is easily carried out and further details are omitted. The resultant Hertz function for the air region is explicitly given by

$$\Pi_z=\frac{ikM}{2}\int_\Gamma F(C)H_0^{(2)}(kS\rho)dC \tag{3.9}$$

where

$$F(C)=\frac{(e^{ikCz}+R_ge^{-ikCz})(e^{ikC(h-z_0)}+R_ie^{-ikC(h-z_0)})}{e^{ikCh}(1-R_gR_ie^{-2ikhC})} \tag{3.10}$$

and

$$R_g=R_g(C)=\frac{N_gC-C_g}{NC_g+C_g} \tag{3.11}$$

$$R_i=R_i(C)=\frac{N_iC-C_i}{N_iC+C_i}. \tag{3.12}$$

It can be immediately noted that the integrand has poles where

$$1-R_g(C)R_i(C)e^{2ikhC}=0.$$

This is the (complex) resonance equation obtained in the previous section by intuitive reasoning.

The integral may be evaluated by using function-theoretic means. The contour is transformed to the S plane. Thus eq (9) becomes

$$\Pi_z=\frac{ikM}{2}\int_\Gamma F(C)H_0^{(2)}(kS\rho)\frac{SdS}{C} \tag{3.13}$$

where the contour Γ may now be taken as the real axis from $-\infty$ to $+\infty$ in the S plane.

The contour is now closed by semicircles in the lower half-plane as indicated in figure 3. Because of the branch point at $S=+1$ and its associated branch line drawn vertically downward, the closing contour runs from one Riemann sheet to the other in the manner indicated. After making two circuits the contour closes on itself. The contours are indented at other branch points in the manner shown for B on the figure. These branch points are located well below the real axis (i.e., imaginary part of $S>>1$) and the corresponding branch cut integrations lead to waves which are heavily damped provided

$$\frac{h}{\rho}<<\left|\frac{1-N_i^2}{N_i}\right| \text{ and } k\rho\,|N_i^2-1|>>1. \tag{3.14}$$

Now the line integral around the complete circuit in the two sheeted Riemann surface is equal to $-4\pi i$ times the sum of the residues of the poles of the integrand. The poles which occur in pairs are located on both Riemann sheets. For highly conducting walls, a number (at least one) is located just below the real axis between the origin and the branch point at $+1$. The remainder are located along or near the negative imaginary axis. The contribution from these latter poles is very small and they correspond to the waveguide modes beyond "cut-off."

The contribution along the semicircles is seen to vanish if the radius R approaches infinity. This is assured by the presence of the Hankel function $H_0^{(2)}(kS\rho)$ which is exponentially decreasing in the lower half-plane of S. Consequently, each of the two integrations along the

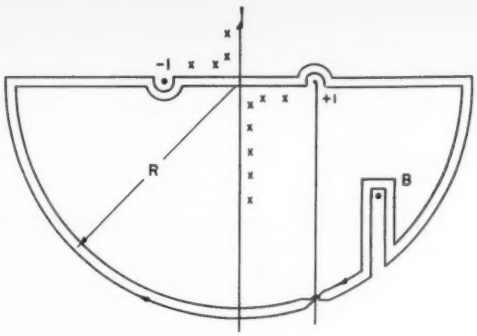


FIGURE 3. *S plane.*
Branch point; x, poles.

real axis are approximately equal to $-2\pi i$ times the sum of the residues at $S=S_n$. When the integral is expressed in the original C plane the residue series may be regarded as the contributions from the poles at $C=C_n$ which for $n=0, 1, 2, \ldots$ are in the first quadrant and for $n=-1, -2, -3, \ldots$ are in the third quadrant. This leads to

$$\Pi_z = \pi k M \sum_{n=-\infty}^{\infty} \left[\frac{1}{\frac{\partial}{\partial C}\frac{1}{F(C)}} \right]_{C=C_n} H_0^{(2)}(kS_n\rho) \tag{3.15}$$

where the square bracket term is the residue of the function $F(C)$ at the pole $C=C_n$. Carrying out the differentiation and making use of the resonance condition

$$R_g(C_n) R_i(C_n) e^{-2ikhC_n} = 1 \tag{3.16}$$

leads readily to

$$\Pi_z \cong \frac{i\pi M}{h} \sum_{n=-\infty}^{+\infty} H_0^{(2)}(kS_n\rho)\, f_n(z_0)\, f_n(z)\, \delta_n(C) \tag{3.17}$$

where

$$\delta_n(C) = \left[1 + i \frac{\partial [R_i(C) R_g(C)]/\partial C}{2khR_i(C)R_g(C)} \right]^{-1}_{C=C_n} \tag{3.18}$$

where

$$f_n(z) = \frac{e^{ikC_n z} + R_g(C_n) e^{-ikC_n z}}{2[R_g(C_n)]^{\frac{1}{2}}} \tag{3.19}$$

and $f_n(z_0)$ has exactly the same form.

When the walls are perfectly conducting $R_i(C)=R_g(C)=1$, the factor $\delta_n(C)$ becomes unity if $n=1, 2, 3, \ldots$ and becomes $\frac{1}{2}$ if $n=0$, and $f_n(z)=\cos kC_n z$. The above expression can then be written

$$\Pi_z = \frac{i\pi M}{h} \sum_{n=0,1,2,\ldots}^{\infty} \epsilon_n H_0^{(2)}(kS_n\rho) \cos(kC_n z_0) \cos(kC_n z) \tag{3.20}$$

where $\epsilon_0=1$, $\epsilon_n=2 (n\neq 0)$. The corresponding value of the electric field component E_z can be expressed

$$E_z = \left(k^2 + \frac{\partial^2}{\partial z^2}\right)\Pi_z = \frac{\mu\omega I ds}{4h} \sum_{n=0,1,2,\ldots}^{\infty} \epsilon_n S_n^2 H_0^{(2)}(kS_n\rho) \cos(kC_n z_0) \cos(kC_n z) \tag{3.21}$$

which is in agreement with eq (2.4) obtained from physical intuitive reasoning.

The extension to the case when the source is a vertical magnetic dipole is simple. Formally the above results are still valid if I is replaced by the magnetic current K. E_z then becomes H_z and the field is essentially horizontally polarized. The reflection coefficients $R_g(C)$ and $R_i(C)$ are to be replaced by their counterparts for horizontal polarization.

Explicitly these are given by

$$R_g^h(C)=\frac{C-N_gC_g}{C+N_gC_g} \tag{3.22}$$

and

$$R_i^h(C)=\frac{C-N_iC_i}{C+N_iC_i}. \tag{3.23}$$

It may be noted, since $R(-C)\ R(C)=1$ where R is any of the reflection coefficients R_i, R_g, R_i^h, or R_g^h, that for negative values of n

$$C_{-1}=-C_0$$

$$C_{-2}=-C_1$$

$$C_{-(n+1)}=-C_n.$$

It then follows that the summations $\sum_{-\infty}^{+\infty}\ldots$ may be replaced by $2\sum_{0}^{\infty}\ldots$ everywhere since $\delta_n(C)=\delta_n(-C)$.

For convenience in numerical computation, it is convenient to express the field components as a ratio to the quantity

$$E_0=i(\eta/\lambda)Ids(e^{-ik\rho})/\rho \qquad (\eta\cong 120\pi). \tag{3.24}$$

E_0 is the field of the source at a distance ρ on a perfectly conducting ground. Thus for both the source and the observer near the ground it is not difficult to show by means of eqs (1), (17), and (19) that

$$E_z=WE_0$$

where

$$W\cong -i\pi\frac{\rho}{h}e^{ik\rho}\sum_{n=0}^{\infty}\delta_nS_n^2H_0^{(2)}(kS_n\rho) \tag{3.25}$$

$$E_\rho=SE_0$$

where

$$S\cong\frac{\pi}{N_g}\frac{\rho}{h}e^{ik\rho}\sum_{n=0}^{\infty}\delta_nS_nH_1^{(2)}(kS_n\rho) \tag{3.26}$$

and

$$H_\phi=TE_0/\eta$$

where

$$T\cong -\pi\frac{\rho}{h}e^{ik\rho}\sum_{n=0}^{\infty}\delta_nS_nH_1^{(2)}(kS_n\rho). \tag{3.27}$$

In the above it has been assumed that $|N_g|^2>>1$.

When $k\rho>>1$, corresponding to the "far-zone," the above expressions may be simplified since the Hankel functions may be replaced by the first term of their asymptotic expansion.[2]

---

[2] The relevant expansions are

$$H_0^{(2)}(x)\cong\left(\frac{2}{\pi x}\right)^{\frac{1}{2}}e^{-ix}e^{i\frac{\pi}{4}}\left[1-\frac{1}{8ix}+\frac{9}{2(8ix)^2}\cdots\right]$$

$$H_1^{(2)}(x)\cong\left(\frac{2}{\pi x}\right)^{\frac{1}{2}}e^{-ix}e^{i\frac{3\pi}{4}}\left[1+\frac{3}{8ix}-\frac{15}{2(8ix)^2}\cdots\right].$$

This leads to the compact result

$$\begin{bmatrix} W \\ S \\ T \end{bmatrix} \cong \frac{(\rho/\lambda)^{\frac{1}{2}}}{(h/\lambda)} e^{i\left[\frac{2\pi\rho}{\lambda}-\frac{\pi}{4}\right]} \sum_{n=0}^{\infty} \delta_n \begin{bmatrix} S_n^{3/2} \\ -S_n^{\frac{1}{2}}/N_g \\ S_n^{\frac{1}{2}} \end{bmatrix} e^{-i2\pi S_n \rho/\lambda} \tag{3.28}$$

which is valid for $\rho >> \lambda$. As expected, the ratio of W to T for a given mode is S_n which for low order or grazing modes is of the order of unity. The ratio of S to T for a given mode is $-1/N_g$ which is very small compared to unity; in fact, it vanishes for a perfectly conducting ground as it must.

3.2. Horizontal Dipole Excitation

The previous section contains the formulation for a vertical dipole source. The corresponding treatment for a horizontal dipole is also quite straightforward although the lack of symmetry increases the complexity. Often in the radio propagation literature the statement is made that the fields of a *horizontal electric* dipole are the same as, or proportional to, the fields of a *vertical magnetic* dipole at the same location. This is only true broadside to the horizontal dipole where the field is purely TE (transverse electric) or horizontally polarized. For other directions, the field has a TM (transverse magnetic) component corresponding to vertical polarization. The modes corresponding to the TM waves may have much smaller attenuation than the modes of the TE type and thus it is desirable to formulate the problem directly with a horizontal dipole source.

As in the previous section the earth and the ionosphere are assumed to be bounded by parallel planes separated by a distance h. Choosing a rectangular coordinate system (x, y, z), the dipole is located at $z=z_0$ and is parallel to the x axis (see fig. 4).

The solution for a horizontal dipole over a homogeneous flat earth with no ionosphere (i.e., $h\to\infty$) was obtained by Sommerfeld many years ago. The generalization for the two interfaces is quite straightforward. A Hertz vector is introduced which has both an x component Π_x and a z component Π_z. The fields in terms of these are

$$\begin{aligned}
E_x &= k^2\Pi_x + \frac{\partial}{\partial x}\left[\frac{\partial \Pi_x}{\partial x} + \frac{\partial \Pi_z}{\partial z}\right], \\
E_y &= \frac{\partial}{\partial y}\left[\frac{\partial \Pi_x}{\partial x} + \frac{\partial \Pi_z}{\partial z}\right], \\
E_z &= k^2\Pi_z + \frac{\partial}{\partial z}\left[\frac{\partial \Pi_x}{\partial x} + \frac{\partial \Pi_z}{\partial z}\right], \\
H_x &= i\epsilon\omega \frac{\partial \Pi_z}{\partial y}, \\
H_y &= -i\epsilon\omega\left[\frac{\partial \Pi_z}{\partial x} - \frac{\partial \Pi_x}{\partial z}\right], \\
H_z &= -i\epsilon\omega \frac{\partial \Pi_x}{\partial y}.
\end{aligned} \tag{3.29}$$

As before a subscript g or i is added to these quantities when specific reference is made to the ground or the ionosphere, respectively.

The boundary conditions at the interfaces $z=0$ and $z=h$ are that tangential components of the fields are continuous. This, in turn, requires that $k^2\Pi_x$, $\partial\Pi_x/\partial x + \partial\Pi_z/\partial z$, $i\epsilon\omega\Pi_z$ and $i\epsilon\omega\partial\Pi_x/\partial z$ are each continuous at these interfaces. Integral representations of Π_x and Π_z which

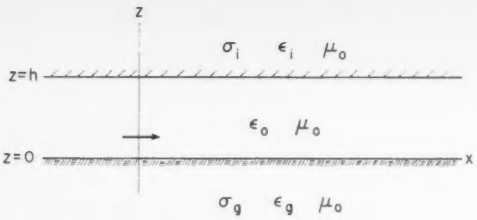


FIGURE 4. *Rectangular coordinate system for the horizontal dipole between two plane interfaces.*

are suitable for matching are

$$\Pi_x = M\frac{e^{-ikR}}{R} + \int_\Gamma [U(C)e^{-ikCz} + V(C)e^{+ikCz}]H_0^{(2)}(kS\rho)dC \tag{3.30}$$

$$\Pi_z = \frac{\partial}{\partial x}\int_\Gamma [X(C)e^{-ikCz} + Y(C)e^{+ikCz}]H_0^{(2)}(kS\rho)dC \tag{3.31}$$

for $0<z<h$. Similar expressions are used for the x and z components of the Hertz vector in the spaces $z<0$ and $z>h$. Applying the boundary conditions involving Π_x only, leads directly to the following solutions for the unknown coefficients in Π_x:

$$U(C) = \left[\frac{R_g^h + R_g^h R_i^h \exp[-2ikC(h-z_0)]}{1 - R_g^h R_i^h \exp(-2ikCh)}\right]\exp(-ikCz_0) \tag{3.32}$$

$$V(C) = \left[\frac{R_i^h + R_g^h R_i^h \exp[-2ikCz_0]}{1 - R_g^h R_i^h \exp(-2ikCh)}\right]\exp[-ikC(2h-z_o)] \tag{3.33}$$

where R_g^h and R_i^h are the complex reflection coefficients given by eqs (22) and (23) and they are also functions of C.

The remaining two boundary conditions, namely, the continuity of $i\epsilon\omega\Pi_z$ and $\partial\Pi_z/\partial z + \partial\Pi_x/\partial x$ enable the coefficients $X(C)$ and $Y(C)$ to be found in terms of $U(C)$ and $V(C)$. The connecting relations are

$$X(C) = \frac{P - QR_g \exp(-ikCh)}{1 - R_i R_g \exp(-2ikCh)} \tag{3.34}$$

and

$$Y(C) = \frac{PR_i \exp(-2ikCh) - Q\exp(-ikCh)}{1 - R_i R_g \exp(-i2kCh)} \tag{3.35}$$

where

$$P = [e^{-ikCz_0} + U + V]\left(\frac{1+R_g}{2ikC}\right) \tag{3.36}$$

and

$$Q = [e^{-ikC(h-z_0)} + Ue^{-ikCh} + Ve^{ikCh}]\left(\frac{1+R_i}{2ikC}\right). \tag{3.37}$$

It is understood that R_i, R_g are functions of C and are defined by eqs (11) and (12).

The integral for Π_x can be observed to have precisely the same form as the z component of a (magnetic) Hertz vector for a vertical magnetic dipole. This in turn has the same general form as the z component of the (electric) Hertz vector for a vertical electric dipole. The residue series representation for Π_x is given by

$$\Pi_x \cong \frac{\pi i M}{h}\sum_m H_0^{(2)}(kS_m\rho) f_m^h(z_0) f_m^h(z)\delta_m^h(C) \tag{3.38}$$

where

$$\delta_m^h(C) = \left[1 + i\frac{dR_i^h(C)R_g^h(C)/dC}{2khC}\right]_{C=C_m}$$

where the summation is over the poles of the integrand at $C=C_m$ of the integral in eq (30). These are solutions of

$$R_1^h(C)R_g^h(C)e^{-2ikhC}=e^{-i2\pi m} \tag{3.39}$$

for integral values of m.

The height functions have the form

$$2f_m^h(z)=[R_g^h(C)]^{-\frac{1}{2}}e^{+ikCz}+[R_g^h(C)]^{\frac{1}{2}}e^{-ikCz}. \tag{3.40}$$

Of particular interest is the vertical magnetic field component. It is given by

$$\begin{aligned} H_z&=-i\epsilon\omega\frac{\partial\Pi_z}{\partial y}=-i\epsilon\omega\frac{\partial\Pi_z}{\partial\rho}\sin\phi \\ &=\sin\phi\,\frac{Idsk}{4h}\sum_m S_mH_1^{(2)}(kS_m\rho)f_m^h(z_0)f_m^h(z)\delta_m^h(C). \end{aligned} \tag{3.41}$$

When $|kS_m\rho|>>1$ or when $\rho>>\lambda$, the first term of the asymptotic expansion of the Hankel function $H_1^{(2)}(kS_m\rho)$ may be employed. This leads to

$$H_z\cong\frac{\sin\phi}{\eta}E_0\frac{(\rho/\lambda)^{\frac{1}{2}}}{(h/\lambda)}\frac{e^{i\pi/4}}{2}\sum_m S_m^{\frac{1}{2}}f_m(z_0)f_m(z)e^{ik(1-S_m)\rho}\delta_m^h(C). \tag{3.42}$$

The other field components involve integrals which may be treated in the same way. Also of great interest is the vertical electric field. It is not difficult to show that

$$E_z=-\frac{\partial}{\partial x}\left[\frac{kM\pi}{h}\sum_n H_0^{(2)}(kS_n\rho)\,g_n(z_0)f_n(z)\,\delta_n(C)\right] \tag{3.43}$$

where

$$2g_n(z_0)=C_n[R_g(C_n)]^{-\frac{1}{2}}\exp(ikC_nz)-C_n[R_g(C_n)]^{\frac{1}{2}}\exp(-ikC_nz). \tag{3.44}$$

It may be noted that

$$kg_n(z_0)=-i\frac{\partial}{\partial z_0}f_n(z_0). \tag{3.45}$$

The summation is now over the roots $C=C_n$ of the equation

$$R_t(C)R_g(C)e^{-i2\pi kCh}=e^{-i2\pi n}. \tag{3.46}$$

When $\rho>>\lambda$ the above expression simplifies to

$$E_z\cong\cos\phi\;E_0\frac{(\rho/\lambda)^{\frac{1}{2}}}{(h/\lambda)}\frac{e^{-i\pi/4}}{2}\sum_n S_n^{\frac{1}{2}}f_n(z)\,g_n(z_0)e^{ik(1-S_n)\rho}\delta_n(C). \tag{3.47}$$

When $|kC_nz_0|\ll1$ and $|kC_nz|\ll1$ the preceding simplifies even further to

$$E_z\cong\cos\phi\;E_0\frac{1}{N_g}\frac{(\rho/\lambda)^{\frac{1}{2}}}{(h/\lambda)}\frac{e^{-i\pi/4}}{2}\sum_n S_n^{\frac{1}{2}}e^{ik(1-S_n)\rho}\delta_n(C). \tag{3.48}$$

4. Properties of the Modes for Flat Earth Case

4.1. Vertical Polarization

Much has been written in the literature on the numerical characteristics of the modes. Controversy concerning the method of numbering the modes has also arisen. It is the opinion that much of this discussion has been unnecessarily involved. The important thing is to sum over all modes which are excited by the dipole. Consequently only these modes need be

considered. Because of the form of the integration contour Γ the relevant solutions must satisfy

$$R_g(C)R_i(C)e^{-2ikhC}=e^{-i2\pi n} \tag{4.1}$$

and have their real and imaginary parts positive. That is, C_n is located in the first quadrant. The numbering is then assigned in such a way that there is continuity in the limiting case of perfect conductivity (i.e., $R_g=R_i=1$).

$$C_n=\frac{\pi n}{kh} \text{ with } n=0,1,2,3, \ldots .$$

As Dr. H. H. Howe points out, this is not quite unambiguous when both the ground and the ionosphere are both imperfectly conducting. The more general statement of the rule is [23]:

For a fixed value of kh, determine n on the assumption of perfectly conducting walls, then σ_g and σ_i in turn are to decrease *continuously* to their prescribed values while C varies continuously. For walls of high but finite conductivity this means that mode 0 has a minimum attenuation and the other modes have successively higher attenuation as n increases. For poorly conducting walls, this is not necessarily so, and in fact, in cases of most practical interest for the vlf band the mode of lowest attenuation is of order one.[3]

Numerical values for C_n are available and will not be quoted here. Some properties of the modes, however, may be simply obtained without resorting to a full numerical solution. For example, if the walls are highly conducting the reflection coefficients may be approximated as follows:

$$R_g(C)=\frac{N_gC-C_g}{N_gC+C_g}\cong 1-\frac{2C_g}{N_gC}\cong \exp\left(-\frac{2}{N_gC}\right) \tag{4.2}$$

$$R_i(C)=\frac{N_iC-C_i}{N_iC+C_i}\cong 1-\frac{2C_i}{N_iC}\cong \exp\left(-\frac{2}{N_iC}\right) \tag{4.3}$$

subject to $|C|^2>>|N_g|^{-2}$ and $|N_i|^{-2}$. Then the resonance equation is simplified to

$$khC=\pi n+i\,\frac{\Delta}{C} \tag{4.4}$$

where $\Delta=(1/N_i+1/N_g)$.

Regarding Δ/C as a small quantity, this can be solved to give

$$S_n\cong\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{\frac{1}{2}}-i\,\frac{\epsilon_n}{2kh}\,\Delta\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{-\frac{1}{2}} \tag{4.5}$$

where $\epsilon_0=1$, $\epsilon_n=2$, $(n\neq 0)$.

The magnitude of the second term must be small compared to the first term for the above perturbation method to be valid. This restriction and the previous one are both met if simultaneously

$$kh|\Delta|<<1 \text{ and } \frac{|\Delta|}{kh}<<1-\left(\frac{\pi n}{kh}\right)^2.$$

Now for highly conducting walls $\sigma_g>>\epsilon_g\omega$ and $\sigma>>\epsilon_i\omega$ and thus

$$\Delta\cong\sqrt{i}\,|\Delta|\cong\sqrt{i}\left[\left(\frac{\epsilon\omega}{\sigma_g}\right)^{\frac{1}{2}}+\left(\frac{\epsilon\omega}{\sigma_i}\right)^{\frac{1}{2}}\right]. \tag{4.6}$$

---

[3] The mode numbering system described above is somewhat different from Budden. For a fixed value of σ_i he starts with a very small value of kh, increases it continuously and requires that C varies continuously for the same n value. [13]

Consequently

$$\text{Re } S_n \cong \left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{\frac{1}{2}} + \frac{\epsilon_n |\Delta|}{2\sqrt{2}kh}\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{-\frac{1}{2}} \tag{4.7}$$

and

$$\text{Im } S_n \cong -\frac{\epsilon_n |\Delta|}{2\sqrt{2}kh}\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{-\frac{1}{2}}. \tag{4.8}$$

The influence of finite conductivity is thus to increase the real part of S_n and consequently the phase velocity $c/\text{Re } S_n$ is decreased relative to the free space value c. As expected the finite conductivity produces damping and the resulting attenuation factor is $-k$ Im S_n in nepers per unit distance, to this approximation.

The above approximate formulas for the real and imaginary parts of S_n are the ones usually encountered. They have been quoted by Schumann [5] for example. It is not also appreciated that they are not applicable for a mode which is near cutoff. This should be evident, however, from the second inequality given above. To relax this restriction, the resonance equation

$$khC = \pi n + i\frac{\Delta}{C} \text{ (valid for } |\Delta| kh << 1) \tag{4.9}$$

is solved as a quadratic in C to yield

$$2C_n = \left(\frac{\pi n}{kh}\right) \pm \left[\left(\frac{\pi n}{kh}\right)^2 + 4i\frac{\Delta}{kh}\right]^{\frac{1}{2}}. \tag{4.10}$$

The positive sign before the radical is chosen since it reduces to $C=(\pi n/kh)$ when Δ approaches zero as it must. The corresponding form for S_n is then given by

$$S_n = \left\{1-\left(\frac{\pi n}{kh}\right)^2 \frac{1}{4}\left[1+\sqrt{1+4i\frac{\Delta kh}{(\pi n)^2}}\right]^2\right\}^{\frac{1}{2}}. \tag{4.11}$$

When $n=0$, this simplifies to

$$S_0 = \left[1-i\frac{\Delta}{kh}\right]^{\frac{1}{2}}$$

which reduces to eq (5) when $n=0$ and $|\Delta| << kh$. Now since $|\Delta| kh << 1$ the radical can be expanded for $n>0$ to yield

$$S_n \cong \left[1-\left(\frac{\pi n}{kh}\right)^2 - i\frac{2\Delta}{kh}\right]^{\frac{1}{2}} \qquad \text{for } n=1,2,3, \ldots. \tag{4.12}$$

The preceding discussion concerns walls which are highly conducting. The approximate solution obtained would indicate that the attenuation increases indefinitely as the conductivity of the walls decreases. Such is true as long as $|\Delta| << 1$. For very poor conductivities this condition becomes violated. When $|\Delta| kh$ is of the order of unity, it is apparently necessary to solve the resonance equation by numerical or graphical means. This approach is described briefly in a later section. As it turns out, for a given value of n, the attenuation reaches a maximum value as $|\Delta|$ is continuously increased and thereafter diminishes and approaches a broad minimum. To illustrate this interesting phenomenon the resonance equation

$$R_g(C) R_i(C) e^{-2ikhC} = e^{-2\pi i n} \tag{4.13}$$

is solved approximately under the condition that

$$|N_i C| << 1 \qquad \text{and } |N_g C| >> 1.$$

Thus

$$R_i(C) \cong -e^{-2N_iC/C_i} \text{ and } R_g(C) \cong e^{-2C_g/N_gC} \tag{4.14}$$

and therefore

$$\frac{C_g}{N_gC}+\frac{N_iC}{C_i}+ikhC=\pi i(n-\tfrac{1}{2}). \tag{4.15}$$

The zero-order approximation is obtained by replacing $R_i(C)$ by -1 and $R_g(C)$ by $+1$. This would yield

$$C=\overline{C}_n=(n-\tfrac{1}{2})\pi/(kh),\ n=1, 2, 3, \ldots \tag{4.16}$$

as mentioned in section 2. For the first-order perturbation

$$C_i=\left(1-\frac{S^2}{N_i^2}\right)^{\frac{1}{2}} \cong \frac{1}{N_i}\,[N_i^2-1+(\overline{C}_n)^2]^{\frac{1}{2}} \tag{4.17}$$

and

$$C_g=\left(1-\frac{S^2}{N_g^2}\right)^{\frac{1}{2}} \cong 1$$

since $|N_gC| >> 1$. With these simplifications it readily follows that

$$C_n \cong \frac{\pi(n-\frac{1}{2})+\dfrac{i}{N_g\overline{C}_n}}{kh-iN_i^2[N_i^2-1+(\overline{C}_n)^2]^{-\frac{1}{2}}} \tag{4.18}$$

and

$$S_n=(1-C_n^2)^{\frac{1}{2}}.$$

When the upper medium is an ionized region, it is convenient to write

$$N_i^2=1-\frac{i}{L}.$$

It may be shown that for vlf, L is approximately real and has a magnitude of the order of unity. Furthermore, for a highly conducting ground

$$1/N_g \cong G^{\frac{1}{2}}e^{i\pi/4}$$

where

$$G=\frac{\epsilon\omega}{\sigma_g}.$$

Then

$$C_n \cong \frac{\pi(n-\frac{1}{2})+e^{+i3\pi/4}G^{\frac{1}{2}}(\overline{C}_n)^{-1}}{kh-i\left(1-\dfrac{i}{L}\right)\left[(\overline{C}_n)^2-\dfrac{i}{L}\right]^{-\frac{1}{2}}}. \tag{4.19}$$

Assuming $(\overline{C}_n)^2 << L$ (which is true for low order modes), and that L is real, the real and imaginary parts of S_n can be written

$$\text{Re } S_n=\overline{S}_n+\frac{1}{2\sqrt{2}\pi(h/\lambda)\overline{S}_n}\left[(\overline{C}_n)^2\left(\sqrt{L}-\frac{1}{\sqrt{L}}\right)+\sqrt{G}\right] \tag{4.20}$$

and

$$\text{Im } S_n=\frac{-1}{2\sqrt{2}\pi(h/\lambda)\overline{S}_n}\left[(\overline{C}_n)^2\left(\sqrt{L}+\frac{1}{\sqrt{L}}\right)+\sqrt{G}\right] \tag{4.21}$$

where

$$\overline{C}_n=\frac{\pi(n-\frac{1}{2})}{kh}=\frac{(n-\frac{1}{2})}{(2h/\lambda)}$$

and

$$\overline{S}_n=[1-(\overline{C}_n)^2]^{\frac{1}{2}}.$$

It may be observed that for a fixed value of h/λ and $\sqrt{G}$ the attenuation factor, $-k$ Im S_n, has a broad minimum when $L=1$. For L somewhat less than unity, the attenuation factor varies as $L^{-\frac{1}{2}}$ or directly as the square root of the effective conductivity of the ionosphere. On the other hand, for L somewhat greater than unity the attenuation factor varies as $L^{\frac{1}{2}}$ or inversely as the square root of the effective conductivity.

The excitation of the modes is proportional to the quantity

$$\delta_n(C_n)=\left[1+i\frac{\partial[R_g(C)R_i(C)]/\partial C}{2khR_g(C)R_i(C)}\right]^{-1}_{C=C_n}. \tag{4.22}$$

When $kh|\Delta|<<1$, where $\Delta=1/N_g+1/N_i$, it follows from eqs (2) and (3) that

$$R_g(C_n)R_i(C_n)\cong\exp\,[-2\Delta/C_n],$$

and

$$[\partial[R_g(C)R_i(C)]/\partial C]_{C=Cn}=\frac{2\Delta}{C_n^2}\exp\,[-2\Delta/C_n]$$

and thus

$$\delta_n(C_n)=\left[1+i\frac{\Delta}{khC_n^2}\right]^{-1}. \tag{4.23}$$

Now the resonance condition states that

$$ikhC_n+\Delta/C_n=i2\pi n \tag{4.24}$$

and for $n=0$ this leads immediately to

$$\delta_0(C_0)=1/2$$

while for $n=1,2,3,\ldots$

$$\delta_n(C_n)=\left(2-\frac{\pi n}{khC_n}\right)^{-1}\cong 1.$$

On the other hand, if the upper medium is very poorly conducting such that

$$|N_iC_n|\ll 1$$

and the lower medium is highly conducting

$$|N_gC_n|\gg 1$$

it follows that

$$R_g(C_n)R_i(C_n)\cong-\exp\left[-\frac{2N_iC_n}{C_i}-\frac{2}{N_gC_n}\right] \tag{4.25}$$

and

$$\left[\frac{\partial}{\partial C}R_g(C)R_i(C)\right]_{C=C_n}\cong-\left[\frac{2N_i}{C_i}-\frac{2}{N_gC_n^2}\right]R_g(C_n)R_i(C_n). \tag{4.26}$$

Thus

$$\delta_n(C_n)\cong\left[1-\frac{i}{kh}\left(\frac{N_i}{C_i}-\frac{1}{N_gC_n^2}\right)\right]^{-1}\cong 1 \tag{4.27}$$

for n=1,2,3, . . . since the term in parentheses is always small compared to unity.

4.2. Horizontal Polarization

In the case when the excitation is by a vertical magnetic dipole or horizontal electric dipole the modes excited may be of a transverse electric (TE) type. The appropriate modal equation is

$$R_g^h(C)\,R_i^h(C)\,e^{-i2khC}=e^{-i2\pi(m-1)} \tag{4.28}$$

where $m=1,2,3,\ldots$.

Now if $|C/N_g|$ and $|C/N_i|\ll 1$

$$R_g^h(C)\cong -\exp(-2C/N_gC_g) \tag{4.29}$$

and

$$R_i^h(C)\cong -\exp(-2C/N_iC_i)\cdot \tag{4.30}$$

The modal equation is thus simplified to

$$C\left[\frac{1}{N_iC_i}+\frac{1}{N_gC_g}\right]+ihkC=i\pi m \tag{4.31}$$

remembering that

$$C_i=\left(1-\frac{S^2}{N_i^2}\right)^{\frac{1}{2}},\qquad C_g=\left(1-\frac{S^2}{N_g^2}\right)^{\frac{1}{2}},$$

and $S^2=1-C^2$.

A first order solution is obtained by replacing S^2 in the expressions for C_i and C_g by the zero order value, e.g.,

$$S^2\cong 1-\left(\frac{\pi m}{kh}\right)^2\cdot \tag{4.32}$$

The approximate solution of the mode equation is then given by

$$C_m\cong\frac{i\pi m}{\Delta_h^m+ikh} \tag{4.33}$$

where

$$\Delta_h^m\cong\left[N_g^2-1+\left(\frac{\pi m}{kh}\right)^2\right]^{-\frac{1}{2}}+\left[N_i^2-1+\left(\frac{\pi m}{kh}\right)^2\right]^{-\frac{1}{2}} \tag{4.34}$$

and $S_m=(1-C_m^2)^{\frac{1}{2}}$. When $|N_i^2|$ and $|N_g^2|\gg 1-\left(\frac{\pi m}{kh}\right)^2$

it is seen that

$$\Delta_h^m\cong\frac{1}{N_g}+\frac{1}{N_i}=\Delta\cdot$$

For $|\Delta_h^m|\ll kh$,

$$C_m\cong\frac{\pi m}{kh}\left(1+i\frac{\Delta}{kh}\right) \tag{4.35}$$

and

$$S_m\cong\left[1-\left(\frac{\pi m}{kh}\right)^2\right]^{\frac{1}{2}}-i\frac{\Delta}{kh}\left(\frac{\pi m}{kh}\right)^2\left[1-\left(\frac{\pi m}{kh}\right)^2\right]^{-\frac{1}{2}} \tag{4.36}$$

which is valid when the modulus of the second term is small compared to the first.

For highly conducting walls

$$\Delta\cong\sqrt{i}\,|\Delta|=\sqrt{i}\left[\left(\frac{\epsilon\omega}{\sigma_g}\right)^{\frac{1}{2}}+\left(\frac{\epsilon\omega}{\sigma_i}\right)^{\frac{1}{2}}\right] \tag{4.37}$$

and therefore

$$\mathrm{Re}\,S_m\cong\left[1-\left(\frac{\pi m}{kh}\right)^2\right]^{\frac{1}{2}}+\frac{|\Delta|}{\sqrt{2}\,kh}\left[1-\left(\frac{\pi m}{kh}\right)^2\right]^{-\frac{1}{2}}\left(\frac{\pi m}{kh}\right)^2 \tag{4.38}$$

and

$$\text{Im } S_m \cong -\frac{|\Delta|}{\sqrt{2}\,kh}\left[1-\left(\frac{\pi m}{kh}\right)^2\right]^{-\frac{1}{2}}\left(\frac{\pi m}{kh}\right)^2. \tag{4.39}$$

In summary, these are valid when

$$|\Delta| \ll kh\left[1-\left(\frac{\pi m}{kh}\right)^2\right]$$

and $\pi m/kh<1$.

It is rather interesting to note that the above expressions for Re S_m and Im S_m are very similar to the corresponding expressions derived for Re S_n and Im S_n in the case of vertical polarization. [For example compare with eqs (7) and (8).] In the present case, of course, there is no zero order mode but apart from this, the perturbation term involving $|\Delta|$ now has an additional factor $[(\pi m/kh)]^2$ which is less than unity if the mode is above "cutoff." Thus, everything else considered equal, the attenuation factor of the TE mode is decreased relative to the TM mode in the earth-ionosphere waveguide with walls assumed to be of high conductivity.

In the earlier notation it was convenient to represent the refractive index in the form

$$N_i^2=1-\frac{i}{L}$$

where L is a real number which may be comparable to or much less than unity. Thus

$$\Delta_h^m \cong \left[\left(\frac{\pi m}{kh}\right)^2-\frac{i}{L}\right]^{-\frac{1}{2}}+\frac{1}{N_g}. \tag{4.40}$$

The corresponding solution for the modal equation is obtained from

$$S_m=\left[1-\left(\frac{\pi m}{kh-i\Delta_h^m}\right)^2\right]^{\frac{1}{2}}. \tag{4.41}$$

When $L<<1$ this reduces to eq (36).

5. Influence of Earth Curvature

The curvature of the earth has been neglected up to this point. The problem is now formulated in terms of spherical coordinates (r, θ, ϕ), with the earth idealized as a homogeneous sphere of radius a, of conductivity σ_g, and dielectric constant ϵ_g. The lower edge of the assumed homogeneous ionosphere is located at $r=a+h$. The source vertical electric dipole is then located at $r=a+z_0$ and the observer is at $r=a+z$ (see fig. 5). In view of the intrinsic spherical symmetry the fields can be represented in terms of a single scalar function, ψ, as follows [26]

$$\begin{aligned} E_r &= \frac{i}{r}\eta\,\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial\psi}{\partial\theta}\right) \\ E_\theta &= -\frac{i}{r}\eta\,\frac{\partial^2}{\partial\theta\partial r}(r\psi) \\ H_\phi &= -k\,\frac{\partial\psi}{\partial\theta} \\ E_\phi &= H_r=H_\theta=0 \end{aligned} \tag{5.1}$$

where $\eta=(\mu/\epsilon)^{\frac{1}{2}}$ and $k^2=\omega^2(\epsilon-i\sigma/\omega)\mu$.

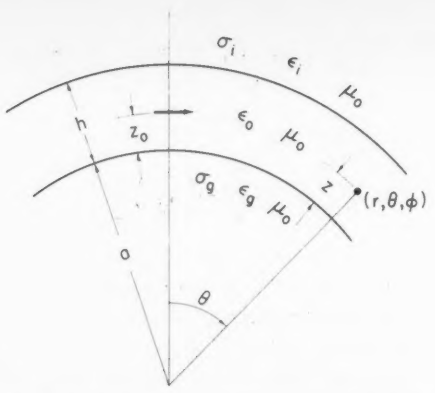


FIGURE 5. *Spherical coordinate system for vertical electric dipole between concentric spherical interfaces.*

As usual, the permeability is taken to be the same as free space for all the regions ($\mu = 4\pi \times 10^{-7}$). A subscript g is affixed to σ, ϵ, etc., when reference is made to the ground and a subscript i for the ionosphere. Since ψ is a solution the wave equation appropriate for the regions, the solution may be represented in terms of spherical wave functions,

$$a_\nu^{(1)} h_\nu^{(1)}(k_g r) P_\nu(-\cos\theta) \qquad \text{for } r<a$$

$$[b_\nu^{(1)} h_\nu^{(1)}(kr) + b_\nu^{(2)} h_\nu^{(2)}(kr)] P_\nu(-\cos\theta) \qquad \text{for } (a+h)>r>a$$

$$c_\nu^{(2)} h_\nu^{(2)}(k_i r) P_\nu(-\cos\theta) \qquad \text{for } r>(a+h).$$

In the above

$$h_\nu^{(1,2)}(kr) = \left(\frac{\pi}{2kr}\right)^{\frac{1}{2}} H_{\nu+\frac{1}{2}}^{(1,2)}(kr) \tag{5.2}$$

where $H_{\nu+\frac{1}{2}}^{(1,2)}(kr)$ is the Hankel function of the first or second kind of order $\nu+\frac{1}{2}$ with argument kr. $P_\nu(-\cos\theta)$ is the hypergeometric series which is a special case of the hypergeometric function $F(\alpha, \beta, \gamma, z)$ namely

$$P_\nu(-\cos\theta) = F\left(-\nu, \nu+1, 1, \frac{1+\cos\theta}{2}\right). \tag{5.3}$$

The reason $P_\nu(-\cos\theta)$ is employed rather than $P_\nu(+\cos\theta)$ is due to the fact that ψ must be regular on a ray $\theta=\pi$, whereas $\theta=0$ is to contain the singularity which is the source of the field. Sommerfeld [26] has pointed out that

$$\lim_{\theta\to 0} F\left(-\nu, \nu+1, 1, \frac{1+\cos\theta}{2}\right) \to \frac{\sin \nu\pi}{\pi} \log \theta^2 \tag{5.4}$$

which illustrates the singular nature of ψ along the polar axis.

The quantity ν is to be found from the boundary conditions that the fields E_θ and H_ϕ are continuous at $r=a$ and $a+h$. This, in turn, requires that $(\eta/r)\,\partial(r\psi)/\partial r$ and $k\psi$ are continuous. Thus, four linear equations in the coefficients $a_\nu^{(1)}$, $b_\nu^{(1)}$, $b_\nu^{(2)}$, and $c_\nu^{(2)}$ are obtained. In order that these yield a nontrivial solution, the four by four determinant of the coefficients should vanish. This requirement is explicitly given by eq (5) as follows:

$$\begin{vmatrix} \left(\frac{\eta_g}{a}\right)\left[\frac{\partial r h_\nu^{(1)}(k_g r)}{\partial r}\right] & -\left(\frac{\eta}{a}\right)\left[\frac{\partial r h_\nu^{(1)}(kr)}{\partial r}\right]_a & -\left(\frac{\eta}{a}\right)\left[\frac{\partial r h_\nu^{(2)}(kr)}{\partial r}\right]_a & 0 \\ k_g h_\nu^{(1)}(k_g a) & -k h_\nu^{(1)}(ka) & -k h_\nu^{(2)}(ka) & 0 \\ 0 & \left(\frac{\eta}{a+h}\right)\left[\frac{\partial r h_\nu^{(1)}(kr)}{\partial r}\right]_{a+h} & \left(\frac{\eta}{a+h}\right)\left[\frac{\partial r h_\nu^{(2)}(kr)}{\partial r}\right]_{a+h} & \left(\frac{\eta_i}{a+h}\right)\left[\frac{\partial r h_\nu^{(2)}(k_i r)}{\partial r}\right]_{a+h} \\ 0 & k h_\nu^{(1)}[k(a+h)] & k h_\nu^{(2)}[k(a+h)] & k_i h_\nu^{(2)}[k_i(a+h)] \end{vmatrix} = 0 \tag{5.5}$$

Such an equation as this was obtained by G. N. Watson in 1919. To solve it for ν without approximation does not seem to be possible, although if the general spherical Hankel functions of complex order and argument could be programed for a computer, an exact numerical solution might be obtained. In view of the idealizations of the model and the uncertainty of the effective electrical constants of the lower ionosphere, however, it does not seem warranted to expend too much effort in this direction. As is so often desirable in physical problems, asymptotic approximations to the rigorous wave functions are introduced which greatly simplify the problem but at the same time lose some generality.

The Debye-Watson representation of the Hankel functions are [26]

$$kr h_{\nu a}^{(1)(2)}(kr) = \left(\frac{\pi kr}{2}\right)^{\frac{1}{2}} H_{\nu+\frac{1}{2}}^{(1)(2)}(kr) \cong \left[1-\frac{(\nu+\frac{1}{2})^2}{(kr)^2}\right]^{-\frac{1}{4}} \exp\left(\mp i\frac{\pi}{4}\right) \exp\left[\pm i\int_{\nu+\frac{1}{2}}^{kr}\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{\frac{1}{2}}dx\right] \tag{5.6}$$

when $|(\nu+\frac{1}{2})/kr|<1$ but not near 1. Also $|\nu+\frac{1}{2}|$ and kr must be large compared to unity. The upper (and lower) signs are to be considered together. This is really a W.K.B. (Wentzel, Kramers, and Brillouin) approximation to the radial part of the wave equation. It is not difficult to show that the resonance equation involving spherical Hankel functions can now be expressed in the equivalent form

$$R_g R_i \exp\left\{-i2\int_{ka}^{k(a+h)}\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]dx\right\} = \exp(-i2\pi n) \tag{5.7}$$

where $n=0, 1, 2, \ldots$.

$$R_g = \frac{k_g\left[1-\frac{(\nu+\frac{1}{2})^2}{(ka)^2}\right]^{\frac{1}{2}} - k\left[1-\frac{(\nu+\frac{1}{2})^2}{(k_g a)^2}\right]^{\frac{1}{2}}}{k_g\left[1-\frac{(\nu+\frac{1}{2})^2}{(ka)^2}\right]^{\frac{1}{2}} + k\left[1-\frac{(\nu+\frac{1}{2})^2}{(k_g a)^2}\right]^{\frac{1}{2}}} \tag{5.8}$$

and

$$R_i = \frac{k_i\left[1-\frac{(\nu+\frac{1}{2})^2}{k^2(a+h)^2}\right]^{\frac{1}{2}} - k\left[1-\frac{(\nu+\frac{1}{2})^2}{k_i^2(a+h)^2}\right]^{\frac{1}{2}}}{k_i\left[1-\frac{(\nu+\frac{1}{2})^2}{k^2(a+h)^2}\right]^{\frac{1}{2}} + k\left[1-\frac{(\nu+\frac{1}{2})^2}{k_i^2(a+h)^2}\right]^{\frac{1}{2}}}. \tag{5.9}$$

The functions R_g and R_i quoted above can readily be identified as Fresnel reflection coefficients for complex angles of incidence $\cos^{-1}C$ and $\cos^{-1}C'$, respectively. Furthermore

$$C=[1-S^2]^{\frac{1}{2}} \qquad \text{where } S=\frac{\nu+\frac{1}{2}}{ka}$$

and

$$C'=[1-(S')^2]^{\frac{1}{2}} \qquad \text{where } S'=\frac{\nu+\frac{1}{2}}{k(a+h)}.$$

The resonance equation can thus be written

$$R_g(C)R_t(C')\ \exp\left[-i2k\int_0^h\left[C^2+\frac{2z}{a}\,S^2\right]^{\frac{1}{2}}dz\right]=e^{-2\pi i n} \tag{5.10}$$

where

$$R_g(C)=\frac{N_gC-C_g}{N_gC+C_g}\qquad \text{with } C_g=\left[1-\frac{S^2}{N_g^2}\right]^{\frac{1}{2}} \tag{5.11}$$

and

$$R_t(C')=\frac{N_tC'-C_t'}{N_tC'+C_t'}\qquad \text{with } C_t'=\left[1-\frac{(S')^2}{N_t^2}\right]^{\frac{1}{2}}. \tag{5.12}$$

It can be seen that, in view of the relation

$$(a+h)\ S'=aS,$$

the resonance equation reduces to its flat earth counterpart as h/a tends to zero. In fact, it appears that if $|C|>>(h/a)^{\frac{1}{3}}\cong 1/10$, the effect of curvature can be disregarded. This condition is violated for most of the numerical results given by Al'pert in the region from 15 to 30 kc. He assumed that the modes could be calculated on the basis of a flat earth in all his work [14, 15].

The resonance equation quoted above for a curved earth is only valid if the W.K.B. or second-order approximations to the spherical wave functions are valid. In a later section the corresponding form of the mode equation based on the Airy or third-order approximation is developed following the work of Rydbeck. It is indicated from this more involved analysis that the second-order approximation is valid if

$$\left(\frac{ka}{2}\right)^{\frac{1}{3}}C<<1.$$

As will be seen, this is met for most cases of practical interest if the frequency is less than about 15 kc.

6. Mode Series for Curved Earth

Following the suggestion of Sommerfeld, the field is written as a sum of modes. Thus

$$\psi(r,\theta)=\sum_n D_\nu z_\nu(kr)P_\nu(-\cos\theta) \tag{6.1}$$

for $a<r<a+h$, where

$$z_\nu(kr)=b_\nu^{(1)}h_\nu^{(1)}(kr)+b_\nu^{(2)}h_\nu^{(2)}(kr). \tag{6.2}$$

The factor D_ν is to be determined by insisting that the function $\psi(r,\theta)$ has the proper behavior at the source. The summation is over all integral values of n and the corresponding (complex) values of ν are obtained from the resonance eq (5.5) as described in the previous section.

Invoking the W.K.B. or second order approximation for the spherical wave functions, it follows that

$$z_\nu(kr)\cong \text{const}\left\{R_g^{-\frac{1}{2}}(C_n)\ \exp\left[+ik\int_0^z(C_n^2+\frac{2z}{a}\,S_n^2)^{\frac{1}{2}}\,dz\right]\right.$$
$$\left.+R_g^{\frac{1}{2}}(C_n)\ \exp\left[-ik\int_0^z(C_n^2+\frac{2z}{a}\,S_n^2)^{\frac{1}{2}}\,dz\right]\right\} \tag{6.3}$$

This can be identified immediately as a combination of a downcoming and an upgoing wave.

The ratio of these two at the earth's surface ($z=0$) is $R_g(C_n)$. An alternate representation is

$$z_\nu(kr) \cong \text{const}\left\{ R_i^{-\frac{1}{2}}(C_n') \quad \text{p}\left[+ik\int_z^h \left(C_n^2+\frac{2z}{a}S_n^2\right)^{\frac{1}{2}} dz\right] \right.$$
$$\left. + R_i^{\frac{1}{2}}(C_n') \exp\left[-ik\int_z^h \left(C_n^2+\frac{2z}{a}S_n^2\right)^{\frac{1}{2}} dz\right]\right\} \quad (6.4)$$

which is a combination of an upgoing wave and a downcoming wave. The ratio of these two at the lower edge of the ionosphere ($z=h$) is $R_i(C_n')$. It should be noted that

$$C_n' \cong \left(C_n^2+\frac{2h}{a}S_n^2\right)^{\frac{1}{2}}$$

since $(a+h)\,S_n'=aS_n$ and $h/a << 1$. The internal consistency of these two representations at $z=0$ and h for $z_\nu(kr)$ can be readily demonstrated from the relation

$$1=R_g^{\frac{1}{2}}(C_n)\,R_i^{\frac{1}{2}}(C_n')\exp\left[-ik\int_0^h \left(C_n^2+\frac{2z}{a}S_n^2\right)^{\frac{1}{2}} dz\right] \quad (6.5)$$

which also indicates that the multiplicative constant is the same for the two representations. In what follows the constant can be absorbed into the factor D_ν.

To study the orthogonality properties of the modes, the following integral is considered

$$I=\int_{ka}^{k(a+h)} z_\nu(\rho)z_\mu(\rho)d\rho \quad (6.6)$$

where ν and μ are two sets of modes. Now quite generally the function $z_\nu(\rho)$ satisfies

$$\rho\frac{d^2}{d\rho^2}(\rho z_\nu)+[\rho^2-\nu(\nu+1)]z_\nu=0 \quad (6.7)$$

and there is a similar relation for z_μ. These two equations are now multiplied by z_μ and z_ν, respectively and integrated over the domain ka to kb, to obtain

$$I=\frac{\rho\left(z_\mu\frac{d}{d\rho}(\rho z_\nu)-z_\nu\frac{d}{d\rho}(\rho z_\mu)\right)\Big|_{ka}^{kb}}{\nu(\nu+1)-\mu(\mu+1)}. \quad (6.8)$$

For the important modes, the right hand side of eq (8) is negligibly small if $\mu\neq\nu$ since the numerator vanishes at the limits ka and kb when R_g and R_i approach ±1. For the important case $\mu=\nu$, a normalization factor is defined by

$$N_\nu=\lim_{\mu\to\nu}\int_{ka}^{kb} z_\mu(\rho)z_\nu(\rho)d\rho$$
$$=\int_{ka}^{kb}[z_\nu(\rho)]^2d\rho\cong\frac{2kh}{\delta_n}, \text{ where } \delta_n=\frac{1}{1+\frac{\sin 2kh\,C_n}{2kh\,C_n}}. \quad (6.9)$$

It should be remarked at this point that the modes are not strictly orthogonal since the right-hand side of eq (8) does not vanish identically although it is small compared to N_ν. As the conductivity of the bounding walls approaches infinity the modes would be completely orthogonal.

Multiplying both sides of eq (1) by $z_\nu(\rho)$ and then integrating with respect to ρ from ka to kb, leads to the following formula for D_ν:

$$D_\nu=\frac{1}{N_\nu P_\nu(-\cos\theta)}\int_{ka}^{kb} z_\nu(\rho)\psi(r,\theta)d\rho. \quad (6.10)$$

To actually evaluate D_ν, it is desirable to let $r \to r_0$ and $\theta \to 0$, in which case $\psi(r,\theta) \to \psi_0(r,\theta)$ where ψ_0 is the primary influence which is singular at $(r_0,0)$. For a vertical electric dipole consisting of an infinitesimal element of length ds and carrying a current, I, it is well known that

$$\psi_0(r,\theta) = \frac{Ids}{4\pi r_0 k} \frac{e^{-ikR}}{R} \tag{6.11}$$

where $R = (r_0^2 + r^2 - 2rr_0 \cos\theta)^{\frac{1}{2}}$.

Following the process suggested by Sommerfeld for the determination of the Green's function for the perfectly conducting sphere, the integration in eq (10) is carried out in the immediate neighborhood of the source. For example, $r = r_0(1+\eta)$ where $-\epsilon < \eta < +\epsilon$, $\epsilon << 1$ and $dr = r_0 d\eta$, $z_\nu(kr) \cong z_\nu(kr_0)$, $e^{-ikR} \cong 1$, while

$$\frac{1}{R} \cong \frac{1}{r_0}\left[(1+\eta)^2 + 1 - 2(1+\eta)\left(1 - \frac{\theta^2}{2}\right)\right]^{-\frac{1}{2}} \cong \frac{1}{r_0(\eta^2+\theta^2)^{\frac{1}{2}}}. \tag{6.12}$$

Therefore

$$N_\nu D_\nu \cong \frac{Ids}{4\pi r_0} z_\nu(kr_0) \lim_{\theta \to 0}\left[\frac{1}{P_\nu(-\cos\theta)} \int_{-\epsilon}^{+\epsilon} \frac{d\eta}{(\eta^2+\theta^2)^{\frac{1}{2}}}\right]. \tag{6.13}$$

Now

$$\lim_{\theta \to 0} [P_\nu(-\cos\theta)] \to \frac{\sin \nu\pi}{\pi} \log \theta^2$$

and

$$\lim_{\theta \to 0} \int_{-\epsilon}^{+\epsilon} \frac{d\eta}{(\eta^2+\theta^2)^{\frac{1}{2}}} = \lim_{\theta \to 0} \left\{ \log\left[\epsilon + \sqrt{\epsilon^2+\theta^2}\right] - \log\left[-\epsilon + \sqrt{\epsilon^2+\theta^2}\right] \right\} \to \log \theta^2. \tag{6.14}$$

It then follows that

$$D_\nu = \frac{i}{2kh} \frac{z_\nu(kr_0)}{\sin \nu\pi} \frac{Ids}{4r_0} \delta_n. \tag{6.15}$$

The final form of the function ψ is thus given by

$$\psi(r,\theta) \cong \frac{Idsi}{2khr_0} \sum_{n=0,1,2,\ldots}^{\infty} \frac{z_\nu(kr_0)}{z_\nu(ka)} \frac{z_\nu(kr)}{z_\nu(ka)} \frac{P_\nu(-\cos\theta)}{\sin \nu\pi} \delta_n \tag{6.16}$$

where the second-order or W.K.B. representations may be used for radial functions $z_\nu(kr)$, etc. As can be seen from eq (4) these can be greatly simplified if $z/a << |C_n^2|$, for then

$$\frac{z_\nu(kr)}{z_\nu(ka)} \cong \frac{e^{ikC_n z} + R_g(C_n) e^{-ikC_n z}}{2[R_g(C_n)]^{\frac{1}{2}}} = f_n(z) \tag{6.17}$$

which is the same height-gain function obtained for the flat earth case. For heights even as great as 10 km and frequencies less than 20 kc, this is an excellent approximation. Similarly,

$$\frac{z_\nu(kr_0)}{z_\nu(ka)} \cong f_n(z_0) \tag{6.18}$$

where $z_0 = r_0 - a$ is the height of the source dipole.

The radial field component is of most practical interest and, for the moment, attention will be confined to it. Since

$$E_r = \frac{i}{r} \eta \frac{1}{\sin\theta} \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial\psi}{\partial\theta}\right) \tag{6.19}$$

and

$$\frac{1}{\sin\theta} \frac{\partial}{\partial\theta}\left[\sin\theta \frac{\partial P_\nu(-\cos\theta)}{\partial\theta}\right] + \nu(\nu+1) P_\nu(-\cos\theta) = 0 \tag{6.20}$$

it follows that

$$E_r \cong \frac{Ids\eta}{2khr_0 r} \sum_{n=0}^{\infty} \delta_n f_n(z_0) f_n(z) \frac{\nu(\nu+1)}{\sin \nu\pi} P_\nu(-\cos\theta) \tag{6.21}$$

with $\nu+\frac{1}{2} \cong kaS_n$. This is the final solution of the problem being valid for the air space between the earth and the ionosphere.[4]

For purposes of computation several simplifications can be made. The asymptotic expansion for the Legendre function, given by

$$P_\nu(-\cos\theta) \cong \left(\frac{2}{\pi\nu \sin\theta}\right)^{\frac{1}{2}} \cos\left[\left(\nu+\frac{1}{2}\right)(\pi-\theta)-\frac{\pi}{4}\right] \tag{6.22}$$

is valid if $|\nu| >> 1$ and θ not near 0 or π. Since the imaginary part of $\nu(\pi-\theta)$ is also large for $\pi-\theta$ greater than about 10° or 20°, it follows that

$$P_\nu(-\cos\theta) \cong \left(\frac{1}{2\pi\nu \sin\theta}\right)^{\frac{1}{2}} \exp\left[i\left(\nu+\frac{1}{2}\right)(\pi-\theta)-i\pi/4\right]. \tag{6.23}$$

Furthermore, the source and observer heights are usually sufficiently low that kh_0C_n and $kh_1C_n \ll 1$ and $r_0 \cong r \cong a$.

The simplified form of the field can now be written

$$E_r = E_0\left[\frac{d/a}{\sin d/a}\right]^{\frac{1}{2}} \frac{(d/\lambda)^{\frac{1}{2}}}{(h/\lambda)} e^{i[2\pi(d/\lambda)-(\pi/4)]} \sum_{n=0}^{\infty} \delta_n S_n^{\frac{3}{2}} e^{-i2\pi S_n (d/\lambda)} \tag{6.24}$$

where $d=a\theta$, the arc length between the source and the observer, λ is the free-space wavelength, and $S_n=(1-C_n^2)^{\frac{1}{2}}$. E_0 is the field of the source at a distance d on a flat perfectly conducting earth. For $d/\lambda \gg 1$,

$$E_0 = i(\eta/\lambda)\, Ids\, (e^{-i2\pi d/\lambda})/d. \tag{6.25}$$

As the radius a of the earth tends to infinity it is immediately evident that the flat earth formula given by eq (2.5) is recovered.

7. Antipodal Effects

The general form for the field in the space $a<r<a+h$ for a vertical dipole source has the form

$$E_r = \frac{Ids\eta}{2kha^2} \sum_{n=0}^{\infty} \delta_n \frac{\nu(\nu+1)}{\sin\nu\pi} P_\nu(-\cos\theta) \tag{7.1}$$

where $\delta_0=\frac{1}{2}$, $\delta_n=1 (n\neq 0)$, and

$$\nu(\nu+1) \cong (\nu+\tfrac{1}{2})^2 \cong kaS_n^2. \tag{7.2}$$

Now as mentioned above when θ is not near 0 or π, the Legendre function may be replaced by the first term of its asymptotic expansion. This result quoted above is valid if

$$\frac{1}{|\nu|} \gg (\pi-\theta) \qquad \text{and} \quad \frac{1}{|\nu|} \gg \theta. \tag{7.3}$$

In this region, the modes are simply proportional to

$$\frac{1}{(\sin\theta)^{\frac{1}{2}}} \cos\left[kaS_n(\pi-\theta)-\frac{\pi}{4}\right] \tag{7.4}$$

---

[4] From an analysis by Pekeris; Phys. Rev. **70**, 518 (1946) it may be shown that

$$\frac{P_\nu(-\cos\theta)}{\sin\nu\pi} \simeq H_0^{(2)}(kS_n\rho) - \frac{1}{12}\left(\frac{\rho}{a}\right)^2 H_2^{(2)}(kS_n\rho)$$

plus terms in $(\rho/a)^4$, $(\rho/a)^6$, etc., where $\nu+\frac{1}{2}=kS_n\rho$. Inserting this in eq (21) leads directly back to eq (3.21) when the curvature correction terms are neglected.

which apart from a constant factor can be identified as the linear combination of two peripheral waves of the form

$$\frac{1}{(\sin\theta)^{\frac{1}{2}}}\, e^{-ikaS_n\theta}$$

and

$$\frac{1}{(\sin\theta)^{\frac{1}{2}}}\, e^{-ikaS_n(2\pi-\theta)} e^{i\pi/2}$$

where $\theta<\pi$.

These waves are traveling in opposing directions along the two respective great circle paths $a\theta$ and $a(2\pi-\theta)$ from the source to the observer. It is noticed that there is a $\pi/2$ phase advance which the wave traveling on the long great circle path picks up as it goes through the pole $\theta=\pi$. The linear combination of these two traveling waves is to form a standing wave pattern whose distance Δ_m between minimums is approximately given by

$$k\Delta_m \mathrm{Re} S_n=\pi \qquad \text{or } \Delta_m=\lambda/(2\ \mathrm{Re}\ S_n)$$

subject to

$$-\mathrm{Im}\ S_n \ll \mathrm{Re}\ S_n.$$

As one approaches the pole $\theta=\pi$, the first term in the asymptotic expansion for the Legendre function is inadequate. A more general form is the asymptotic series [27]

$$P_\nu(-\cos\theta)\cong\frac{2}{\pi^{\frac{1}{2}}}\frac{\Gamma(\nu+1)}{\Gamma(\nu+3/2)}\sum_l \frac{[(1/2)l]^2}{(\nu+3/2)_l\, l!}\,\frac{\cos\left[\left(\nu+\frac{2l+1}{2}\right)(\pi-\theta)-\left(\frac{2l+1}{4}\right)\pi\right]}{(2\sin\theta)^{l+\frac{1}{2}}} \tag{7.5}$$

where

$$(\alpha)_l=\alpha(\alpha+1)\ (\alpha+2)\ .\ .\ .\ (\alpha+l-1),$$

for example, $\alpha_0=1$, $\alpha_1=\alpha$, $\alpha_2=\alpha(\alpha+1)$, $\alpha_3=\alpha(\alpha+1)\ (\alpha+2)$, etc.

Since $|\nu|\gg 1$ the factorial functions may be replaced by the first two terms of their asymptotic expansions; this leads to

$$\frac{\Gamma(\nu+1)}{\Gamma(\nu+3/2)}\sim\frac{1}{\nu^{\frac{1}{2}}}\left(1-\frac{3}{8\nu}+0\left(\frac{1}{\nu^2}\right)\right). \tag{7.6}$$

The preceding asymptotic expansion for $P_\nu(-\cos\theta)$ is not usable at and in the vicinity of the pole $\theta=\pi$. In this region a suitable representation is given by [27]

$$P_\nu(-\cos\theta)=J_0(\eta)+\sin^2\left(\frac{\pi-\theta}{2}\right)\left[\frac{J_1(\eta)}{2\eta}-J_2(\eta)+\frac{\eta}{6}J_3(\eta)\right]+0\left(\sin^4\left(\frac{\pi-\theta}{2}\right)\right) \tag{7.7}$$

where $\eta=(2\nu+1)\sin[(\pi-\theta)/2]$. $J_m(\eta)$, for $m=0$, 1, 2, and 3, is the Bessel function of first type of argument η and order m. When $\pi-\theta$ is small the first term is usually sufficient and furthermore

$$\eta\cong(\nu+\tfrac{1}{2})\ (\pi-\theta)\cong kaS_n(\pi-\theta).$$

Thus for mode n, the field in the neighborhood of the pole is proportional to the Bessel function

$$J_0[kaS_n(\pi-\theta)].$$

It is then not surprising to see that the first term of the asymptotic expansion of J_0 is the same as that of $P_\nu(-\cos\theta)$.

8. Resonator Type Oscillations Between Earth and the Ionosphere

At extremely low frequencies (elf), where the wavelength is large compared to the height of the ionospheric reflecting layer, the electric field is essentially radial and only one waveguide

type mode is significant. The field is thus expressed by the first term of the mode series which reads

$$E_r=\frac{Ids\eta}{4kha^2}\frac{\nu(\nu+1)}{\sin \nu\pi}P_\nu(-\cos\theta) \tag{8.1}$$

where $\nu+\frac{1}{2}\cong kaS_0$ and

$$S_0\cong 1-\frac{i}{2kh}\left(\frac{1}{N_i}+\frac{1}{N_g}\right) \tag{8.2}$$

in terms of the relative refractive indices N_i and N_g of the homogeneous ionosphere and the homogeneous ground, respectively. Now at elf $|N_g|>>|N_i|$ and furthermore,

$$N_i\cong\left(\frac{\sigma_i+i\omega\epsilon_i}{i\omega\epsilon}\right)^{\frac{1}{2}}\cong\left(\frac{\sigma_i}{i\omega\epsilon}\right)^{\frac{1}{2}}.$$

Thus

$$S_0\cong 1-\frac{i^{3/2}}{2(\sigma_i\mu\omega)^{\frac{1}{2}}h}.$$

Now as mentioned in section 7, the factor $P_\nu(-\cos\theta)$ may be replaced by an asymptotic expansion if $ka\theta$ or $ka(\pi-\theta)$ is somewhat greater than unity. The field in this case may be regarded as two azimuthal-type traveling waves. Furthermore at the pole (θ near π) where the second of these restrictions is violated, it is possible to use an equivalent representation which correctly accounts for the axial focusing. An alternate viewpoint which is suitable at elf is to consider the field as a superposition of cavity-resonator type modes. It is expected that such a representation would be very good when the circumference of the earth is becoming comparable to the wavelength. A suggestion of this kind was apparently first put forth by Schumann [28].

The starting point is the expansion formula

$$\frac{P_\nu(-x)}{\sin\nu\pi}=-\frac{1}{\pi}\sum_{n=0}^{\infty}P_n(x)\frac{2n+1}{n(n+1)-\nu(\nu+1)} \tag{8.3}$$

where the summation is over integral values of n. This result follows directly from a formula given by Magnus and Oberhettinger [27] (p. 57) which is valid for $\nu\neq 0, \pm 1, \pm 2, \ldots$, and $0\leqq\theta<\pi$. The electric field, for $h/a<<1$, is thus written

$$E_r=\frac{I\,ds\,\nu(\nu+1)}{4\pi a^2\epsilon\omega h}\sum_{n=0}^{\infty}P_n(x)\frac{2n+1}{n(n+1)-\nu(\nu+1)} \tag{8.4}$$

where $x=\cos\theta$. The early terms of the series are then proportional to

$$\begin{aligned}P_0(x)&=1,\\ P_1(x)&=\cos\theta,\\ P_2(x)&=\tfrac{1}{2}(3\cos^2\theta-1),\end{aligned} \tag{8.5}$$

and so on. The configuration of the electric field in the first three cavity modes is depicted in figure 6.

Retaining just the first term it is seen that

$$E_r^\circ=E_r]_{n=0}=\frac{Ids}{4\pi a^2\epsilon h}\frac{1}{i\omega} \tag{8.6}$$

which is independent of θ. Clearly this corresponds to a concentric spherical capacitor energized by a current Ids/h resulting in a constant voltage hE_r° between the plates. On rewriting eq (6) in the form

$$hE_r^\circ=\frac{(Ids/h)}{i\omega C_e} \tag{8.7}$$

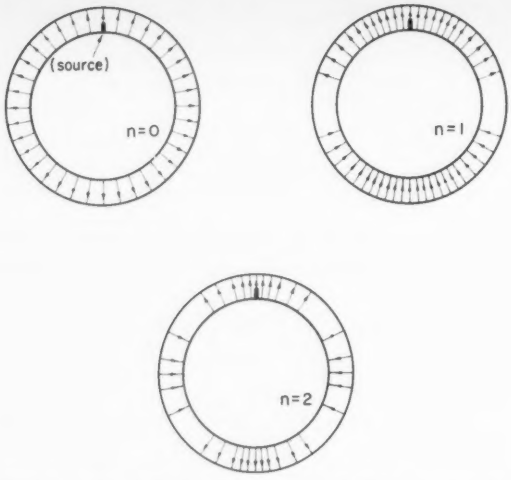


FIGURE 6. *Depicting electric field lines in the first three cavity resonator modes.*

it is seen that

$$C_e=\frac{4\pi a^2\epsilon}{h}$$

which can be identified as the capacity between the spherical surfaces whose areas are both $4\pi a^2$ within the approximation $h/a<<1$.

It has often been suggested that the omnipresent constant voltage gradient in the atmosphere results from the accumulated action of lightning strokes which impart a charge to this earth-ionosphere condenser. For example, when a current surge flows say for 10^{-3} sec with an average of 10^3 amp with an average column height of 3 km (i.e., $ds\cong 2\times 10^3$, then for $h\cong 70$ km it readily follows that

$$hE_r^\circ\cong 1.3 \text{ v} \qquad \text{or } E_r^\circ\cong 2\times 10^{-5} \text{ v/m}.$$

Presumably, many such charges are required to build the field up to its observed value.

Of somewhat more interest are the cavity-resonator oscillations which may be excited. Using the notation of the operational calculus $i\omega$ is formally replaced by p then

$$\nu(\nu+1)\cong -p^2-p^{\frac{3}{2}}\alpha \tag{8.8}$$

where $\alpha=1/[h(\sigma_i\mu)^{1/2}]$. The source dipole moment Ids is in general a function of time. For purposes of illustration, consider

$$Ids=(Ids)_0\, u(t) \tag{8.9}$$

where $u(t)$ is the unit step function at $t=0$. The Laplace transform of the source moment is thus given by

$$\int_0^\infty Ids e^{-pt}dt=\frac{(Ids)_0}{p}. \tag{8.10}$$

The Laplace transform of the field is given by

$$E_r(p)=\frac{(Ids)_0}{h^2C_e}\,(p+\alpha p^{\frac{1}{2}})\sum_{n=0}^{\infty}P_n(x)\,\frac{2n+1}{\omega_n^2+p^2+\alpha p^{\frac{3}{2}}} \tag{8.11}$$

where $\omega_n^2=(a/c)^2\, n(n+1)$. The actual time response of the electric field is denoted $e_r(t)$ and is zero for $t<0$. It is related to $E_r(p)$ by

$$E_r(p)=\int_0^\infty e_r(t)e^{-pt}dt. \tag{8.12}$$

The inversion of this integral equation is a standard problem in operational calculus and has been carried out explicitly by Schumann [27] for a transform which has the form of eq (11). In the present discussion a much simpler approach is used which is justified when the damping is small. It should be noted that $\alpha p^{\frac{1}{2}}$ has already been assumed small compared to p, thus a perturbation method is in order.

When $\alpha=0$ corresponding to no dissipation (i.e., perfectly conducting boundaries)

$$E_r^{(n)}(p)=\frac{(I\,ds)_0}{h^2C_e}\,P_n(x)\,\frac{(2n+1)p}{\omega_n^2+p^2}. \tag{8.13}$$

The poles in the p plane are thus at $p=\pm i\omega_n$. The inversion to the time domain gives

$$e_r^{(n)}(t)=\frac{(I\,ds)_0}{h^2C_e}\,P_n(x)\,(2n+1)\,\cos\,\omega_n t \tag{8.14}$$

which may be verified by noting that the above expressions for $E_r^{(n)}(p)$ and $e_r^{(n)}(t)$ satisfy eq (12).

A step-function dipole source thus excites the static field (i.e., $\omega_0=0$) and the cavity-resonator modes ($n=1, 2, 3, \ldots$). For $a=6{,}400$ km

$$\omega_1/2\pi=10.6 \text{ cps}$$

$$\omega_2/2\pi=18.3 \text{ cps}$$

$$\omega_3/2\pi=25.9 \text{ cps}.$$

To account for finite conductivity it is necessary to solve the equation

$$p^2+p^{3/2}\alpha+\omega_n^2=0 \tag{8.15}$$

which gives the poles for the function $E_r(p)$ in the case when $\alpha\neq 0$. Remembering that $\alpha p^{\frac{1}{2}} << p$, it readily follows that

$$p\cong i\omega_n\left[1+\frac{\alpha}{2\omega_n^{1/2}}e^{i3\pi/4}\right]\cong i\omega_n'-\Omega_n \tag{8.16}$$

where $\omega_n'=\omega_n\left(1-\frac{\alpha}{2^{3/2}\omega_n{}^{1/2}}\right)$ is the resonant frequency and

$$\Omega_n=\frac{\alpha\omega_n^{1/2}}{2^{3/2}}$$

is the damping coefficient. It then easily follows that $\cos\,\omega_n t$ is to be replaced by

$$e^{-\Omega_n t}\,\cos\,\omega_n' t.$$

To this approximation, the effect of finite conductivity is to exponentially damp the oscillations with time. For $a=6{,}400$ km, $h\sim 100$ km, $\sigma_i\sim 10^{-4}$ mhos/m, the time constant is given by

$$\frac{1}{\Omega_n}\cong\frac{4}{\sqrt{n(n+1)}}\text{ sec }(n=1,2,3,\ldots) \tag{8.17}$$

which is rather interesting.

The total field is thus given by

$$e_r(t) \cong \frac{(I\,ds)_0}{h^2 C_e} \sum_{n=0}^{\infty} P_n(x)\,(2n+1) e^{-\Omega_n' t} \cos \omega_n' t \tag{8.18}$$

which is valid for $\alpha^2 t >> 1$ or $t >> 1/(h^2 \sigma_1 \mu)$.

9. Excitation by Horizontal Dipoles for the Curved Earth

The formulation of the theory for a horizontal dipole is similar to that for a vertical dipole. The complexity of the equations is greater, however, because of the nonsymmetry of the problem. Schumann [6] uses this approach in his analysis but his results are not complete as discussed below. The deficiency arises when the eigenfunction series is matched to the source singularity. In the case of the vertical dipole as outlined in the previous sections, this process is relatively straightforward but in the case of the horizontal dipole there is coupling between TE and TM modes which apparently is not accounted for using this technique. An alternative is to set up the problem in terms of a harmonic series representation wherein the summation is over integral values of n, the index of the spherical wave functions. This series is poorly convergent, however, and the Watson technique must be used to transform it to a series of residues at the complex poles ν. Such a procedure was used by Wait [29] for a horizontal dipole over an earth with a homogeneous atmosphere. It would not be difficult to generalize these results to include the influence of the ionospheric reflecting layer. In the present work, however, it seems more instructive to use a different method which makes use of the reciprocity theorem and the results for vertical electric and vertical magnetic dipoles.

For the first part of the problem a vertical magnetic dipole of moment $K\,ds$ is considered. It is located at $r=r_0$ on the polar axis. Due again to the intrinsic symmetry of the problem the fields can be obtained from a single scalar function ψ^h as follows:

$$\begin{aligned} H_r &= \frac{i}{r\eta} \frac{1}{\sin\theta} \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial \psi^h}{\partial\theta}\right) \\ H_\theta &= -\frac{i}{r\eta} \frac{\partial^2}{\partial\theta\,\partial r}(r\psi^h) \\ E_\phi &= k \frac{\partial \psi^h}{\partial\theta} \end{aligned} \tag{9.1}$$

and $H_\phi = E_r = H_\theta = 0$. Such fields are purely of the TE type whereas they were of the TM type for a vertical electric dipole source.

The solution precedes in the same manner as for the vertical electric dipole. Now, however, the boundary conditions are that $(1/\eta r)\partial(r\psi^h)/\partial r$ and $k\psi^h$ are continuous at the concentric spherical interfaces. In a form suitable for application to vlf propagation, the final result is given by [5]

$$\psi^h = \frac{K\,ds}{2khr_0} \sum_{m=1,2,3,\ldots}^{\infty} \frac{\delta_m^h f_m^h(z_0) f_m^h(z)}{\sin \bar{\nu}\pi} P_{\bar{\nu}}(-\cos\theta) \tag{9.2}$$

where

$$\delta_m^h \cong \frac{1}{1 - \dfrac{\sin 2khC_m}{2khC_m}}, \tag{9.3}$$

$$2f_m^h(z) = [R_g^h(C_m)]^{-\frac{1}{2}} \exp\,[ikC_m z] + [R_g^h(C_m)]^{\frac{1}{2}} \exp\,[-ikC_m z] \tag{9.4}$$

[5] To conform with standard waveguide practice, the TE mode of lowest attenuation is denoted $m=1$.

and similarly for $f_m^h(z_0)$. The modal equation has the form

$$R_g^h(C)R_i^h(C')\exp\left\{-i2k\int_0^h\left[C^2+\frac{2z}{a}S^2\right]^{\frac{1}{2}}dz\right\}=e^{-2\pi im} \tag{9.5}$$

where

$$R_g^h(C)=\frac{C-N_gC_g}{C+N_gC_g} \qquad \text{with } C_g=\left[1-\frac{S^2}{N_g^2}\right]^{\frac{1}{2}}$$

and

$$R_i^h(C')=\frac{C'-N_iC_i'}{C'+N_iC_i'} \text{ with } C_i'=\left[1-\frac{(S')^2}{N_i^2}\right]^{\frac{1}{2}}.$$

The electric field component E_ϕ is thus given by

$$E_\phi=\frac{Kds}{2hr_0}\sum_m \delta_m^h \frac{f_m^h(z_0)f_m^h(z)}{\sin \bar{\nu}\pi}\frac{\partial P_{\bar{\nu}}(-\cos\theta)}{\partial\theta}. \tag{9.6}$$

Now when the source is a small loop of area da carrying an average circulating current I it follows that

$$Kds=i\mu\omega Ida. \tag{9.7}$$

Furthermore, if the receiving antenna is a horizontal electric antenna of effective length dl, the voltage at the terminals is given by

$$v=E_\phi dl\sin\phi \tag{9.8}$$

where ϕ is the angle subtended by the receiving antenna and the arc joining the two antennas (see fig. 7). Thus the mutual impedance Z_m between the source loop of area da and

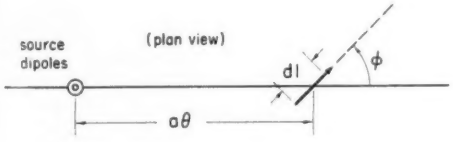


FIGURE 7. *Source vertical magnetic and electric dipoles and receiving horizontal electric dipole for mutual impedance calculation.*

the horizontal receiving antenna of length dl is

$$Z_m=\frac{v}{I}=\frac{i\mu\omega\, da\, dl\sin\phi}{2hr_0}\sum_m\cdots \tag{9.9}$$

where the summand is the same as in eq (6).

The mutual impedance Z_e between a vertical electric dipole source at $\theta=0$, $r=r_0$, and the horizontal receiving antenna is also required. This may be obtained from the scalar function ψ previously obtained. In particular

$$E_\theta=-\frac{i}{r}\eta\frac{\partial^2}{\partial\theta\partial r}(r\psi)\cong -i\eta\frac{\partial^2}{\partial\theta\partial z}\psi \tag{9.10}$$

$$\cong\eta\frac{Ids}{2hr_0}\sum_n\delta_n\frac{f_n(z_0)g_n(z)}{\sin\nu\pi}\frac{\partial}{\partial\theta}P_\nu(-\cos\theta) \tag{9.11}$$

where

$$2g_n(z)=\left[C_n^2+\frac{2z}{a}S_n^2\right]^{\frac{1}{2}}\left\{R_g^{-1/2}\exp\left[+ik\int_0^z\left(C_n^2+\frac{2z}{a}S_n^2\right)^{\frac{1}{2}}dz\right]\right.$$

$$\left.-R_g^{1/2}\exp\left[-ik\int_0^z\left(C_n^2+\frac{2z}{a}S_n^2\right)^{\frac{1}{2}}\right]dz\right\}. \quad (9.12)$$

When $z/a<<|C_n^2|$

$$2g_n(z)\cong C_n[R_g^{-1/2}e^{ikC_nz}-R_g^{1/2}e^{-ikC_nz}]. \quad (9.13)$$

Furthermore if $|kC_nz|<<1$ which is the usual case

$$g_n(z)\cong\Delta_n f_n(z) \quad (9.14)$$

where

$$\Delta_n=\frac{1}{N_g}\left(1-\frac{S_n^2}{N_g^2}\right)^{\frac{1}{2}}\cong\frac{1}{N_g}.$$

The mutual impedance Z_e between the vertical electric dipole and the horizontal receiving antenna dl is thus given by

$$Z_e\cong\frac{\eta\,ds\,dl\cos\phi}{2hr_0}\sum_n\cdots \quad (9.15)$$

where the summand is the same as in eq (11).

It is now a simple matter to write down the field expressions when the source is a horizontal electric antenna carrying a current I of length dl. The antenna or dipole now is considered to be located at $r=r_0$ and $\theta=0$ and oriented in the direction $\phi=0$. The vertical magnetic field at (r, θ, ϕ) is obtained from the relation

$$i\mu\omega H_r da=IZ_m \quad (9.16)$$

which relates the total magnetic flux in a small loop of area da at (r, θ, ϕ) and the vertical magnetic field at the same point. Using eq (9.) it is seen that

$$H_r=\frac{Idl}{2hr}\sum_m\delta_m^h\frac{f_m^h(z_0)f_m^h(z)}{\sin\bar{\nu}\pi}\frac{\partial P_{\bar{\nu}}(-\cos\theta)}{\partial\theta}\sin\phi. \quad (9.17)$$

In a similar fashion, the vertical electric field at (r, θ, ϕ) is obtained from the relation

$$E_r ds=IZ_e \quad (9.18)$$

which relates the voltage in the small vertical antenna of length ds at (r, θ, ϕ) and the vertical electric field at the same point. Thus

$$E_r=\eta\frac{Idl}{2hr}\sum_n\frac{\delta_n f_n(z)g_n(z_0)}{\sin\nu\pi}\frac{\partial}{\partial\theta}P_\nu(-\cos\theta)\cos\phi. \quad (9.19)$$

The other field components can be found from the above expressions for E_r and H_r. Quite generally the field components in spherical coordinates (see fig. 8) can be written in terms of a set of purely TE and TM modes derivable from scalar functions U and V. Thus

$$E_r=\left[\frac{\partial^2(Ur)}{\partial r^2}+k^2rU\right] \quad (9.20)$$

$$E_\theta=\frac{1}{r}\frac{\partial^2}{\partial\theta\partial r}(rU)-\frac{i\mu\omega}{r\sin\theta}\frac{\partial(rV)}{\partial\phi} \quad (9.21)$$

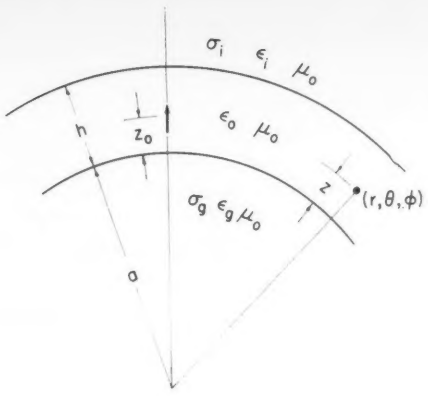


FIGURE 8. *Spherical coordinate system for horizontal electric dipole between concentric spherical interfaces.*

$$E_\phi=\frac{1}{r\sin\theta}\frac{\partial^2}{\partial\phi\partial r}(rU)+\frac{i\mu\omega}{r}\frac{\partial(rV)}{\partial\theta} \tag{9.22}$$

$$H_r=\frac{\partial^2(rV)}{\partial r^2}+k^2rV \tag{9.23}$$

$$H_\theta=\frac{1}{r}\frac{\partial^2}{\partial r\partial\theta}(rV)+\frac{i\epsilon\omega}{r\sin\theta}\frac{\partial}{\partial\phi}(rU) \tag{9.24}$$

$$H_\phi=\frac{1}{r\sin\theta}\frac{\partial^2}{\partial r\partial\phi}(rU)-\frac{i\epsilon\omega}{r}\frac{\partial}{\partial\theta}(rU). \tag{9.25}$$

Since U and V satisfy the equations

$$(\nabla^2+k^2)\begin{matrix}U\\V\end{matrix}=0 \tag{9.26}$$

in the space $a<r<a+h$ they must be made up of solutions of the form

$$\begin{matrix}h_\nu^{(1)}(kr)\\h_\nu^{(2)}(kr)\end{matrix}\quad P_\nu^q(\cos\theta)\quad\begin{matrix}\cos q\phi\\\sin q\phi\end{matrix} \tag{9.27}$$

where q is an integer. Since the field for E_r and H_r has already been prescribed, $q=1$. With some consideration it is seen that

$$U=\sum_n A_n f_n(z)\frac{\partial P_\nu(-\cos\theta)}{\partial\theta}\cos\phi \tag{9.28}$$

and

$$V=\sum_m B_m f_m^h(z)\frac{\partial P_\nu(-\cos\theta)}{\partial\theta}\sin\phi. \tag{9.29}$$

Further, on noting

$$\left[k^2+\frac{\partial^2}{\partial r^2}-\frac{\nu(\nu+1)}{r^2}\right][rh_\nu^{(1,2)}(kr)]=0 \tag{9.30}$$

it is seen that

$$E_r=\sum_n\frac{\nu(\nu+1)}{r}A_n f_n(z)\frac{\partial P_\nu(-\cos\theta)}{\partial\theta}\cos\phi \tag{9.31}$$

and

$$H_r=\sum_m\frac{\bar\nu(\bar\nu+1)}{r}B_m f_m^h(z)\frac{\partial P_{\bar\nu}(-\cos\theta)}{\partial\theta}\sin\phi. \tag{9.32}$$

On comparing eqs (30) and (31) with (17) and (19), it follows that

$$A_n=\frac{\eta Idl}{2h\nu(\nu+1)}\frac{\delta_n g_n(z_0)}{\sin \nu\pi} \tag{9.33}$$

and

$$B_m=\frac{Idl}{2h\bar{\nu}(\bar{\nu}+1)}\frac{\delta_m^h f_m^h(z_0)}{\sin \bar{\nu}\pi}. \tag{9.34}$$

When $|\nu|>>1$ and θ is not near 0 or π the following asymptotic expansion is valid:

$$\frac{1}{\sin \nu\pi}\frac{\partial}{\partial\theta}P_\nu(-\cos\theta)\cong\frac{2e^{-i\pi/4}}{(\sin\theta)^{\frac{1}{2}}}\left(\frac{a}{\lambda}\right)^{\frac{1}{2}}S_n^{1/2}e^{-ika\theta S_n} \tag{9.35}$$

where use has been made of the relation

$$\nu\cong kaS_n.$$

The height function $g(z_0)$ occurring in the expression for E_r can be simplified at low heights. For example, if $|kC_nz_0|<<1$ which is the usual case

$$g_n(z_0)\cong\frac{1}{iN_g}\left(1-\frac{S_n^2}{N_g^2}\right)^{\frac{1}{2}}. \tag{9.36}$$

Thus the vertical electric field of a horizontal dipole is well approximated by

$$E_r\cong E_0\frac{\cos\phi}{N_g}\left[\frac{d/a}{\sin d/a}\right]^{\frac{1}{2}}\frac{(d/\lambda)^{\frac{1}{2}}}{(h/\lambda)}e^{i[(2\pi d/\lambda)-(\pi/4)]}\sum_{n=0}^{\infty}\delta_nS_n^{1/2}e^{-i2\pi S_n(d/\lambda)} \tag{9.37}$$

where

$$E_0=i(\eta/\lambda)Ids(e^{-i2\pi d/\lambda})/d. \tag{9.38}$$

It is of interest to compare this with eq (6.24) for the vertical electric field of a vertical dipole with the same moment. It is seen for a given mode

$$\frac{E_r^{(n)}\text{ (for horizontal dipole)}}{E_r^{(n)}\text{ (for vertical dipole)}}\cong\frac{\cos\phi}{S_nN_g}\left(1-\frac{S_n^2}{N_g^2}\right)^{\frac{1}{2}}. \tag{9.39}$$

Since $S_n\cong 1$, it is seen that the ratio does not depend critically on mode number n, thus

$$\frac{E_r\text{ (for horizontal dipole)}}{E_r\text{ (for vertical dipole)}}\cong\frac{\cos\phi}{N_g}\left(1-\frac{1}{N_g^2}\right)^{\frac{1}{2}}. \tag{9.40}$$

In most cases $|N_g|>>1$ so the ratio is of the order of $1/N_g$ which is small. In particular, at very low frequencies

$$\frac{\cos\phi}{N_g}\cong\left(\frac{\epsilon\omega}{\sigma_g}\right)^{\frac{1}{2}}e^{i\pi/4}\cos\phi \tag{9.41}$$

which indicates that the ratio varies as the square root of the frequency. This is in disagreement with the results of Schumann who finds that dependence is the inverse first power of frequency. In the direction $\phi=0$, of course, the ratio derived here turns out to be nothing more than the "wave-tilt" for a vertically polarized plane wave at grazing incidence on a flat earth. Thus, Schumann's results [7] for the horizontal dipole would seem to be in error.

The horizontal dipole also, of course, radiates horizontal polarization. The simplified expression for H_r can be written by employing the single term asymptotic representation described above. Thus

$$H_r\cong\frac{E_0\sin\phi}{\eta}\frac{(d/\lambda)^{\frac{1}{2}}}{(h/\lambda)}\left[\frac{d/a}{\sin d/a}\right]^{\frac{1}{2}}e^{i[(2\pi d/\lambda)-(\pi/4)]}\sum_{m=1}^{\infty}\delta_m^he^{-ika\theta S_m}S_m^{\frac{1}{2}}f_m^h(z_0)f_m^h(z). \tag{9.42}$$

10. Influence of the Earth's Magnetic Field

In the preceding analysis the earth's magnetic field has tacitly been neglected. To indicate its effect the reflection coefficient for sharply bounded ionosphere with the magnetic field included shall be discussed. The derivation of the general formulas is due to Budden [30] but for highly oblique incidence great simplifications to his results can be made.

The starting point is the magneto-ionic formula of Appleton and Hartree for the complex refractive index μ for a homogeneous ionized medium with superimposed magnetic field. In the region from 70 to 90 km in the ionosphere where very low frequencies are reflected, it is often permissible to employ the quasi-longitudinal approximation of Booker. It is now implied that the waves after they are transmitted into the ionosphere are steeply refracted toward the vertical. Essentially this means that the refractive index does not depend to any great extent on the direction of propagation for temperate and polar latitudes so that

$$\mu^2 \cong 1 - i(\omega_r/\omega) \exp(\pm i\tau) \tag{10.1}$$

where

$$\tan \tau = \omega_L/\nu$$

and

$$\omega_r = \omega_0^2(\nu^2 + \omega_L{}^2)^{-\frac{1}{2}}.$$

In the above

$\omega_0^2 = Ne^2/\epsilon m$,
N = number of electrons per meter^3,
e and m = charge and mass of electrons,
$\epsilon = 8.854 \times 10^{-12}$,
ν = collision frequency,
$\omega_L = (4\pi \times 10^{-7})He/m$, and
H = effective strength of the earth's magnetic field, (i.e., the longitudinal component for propagation in the ionosphere).

It is now desirable to consider four reflection coefficients ${}_{||}R_{||}$, ${}_{||}R_{\perp}$, ${}_{\perp}R_{||}$, and ${}_{\perp}R_{\perp}$ to indicate the complex ratio of a specified electric field in the wave after reflection to a specified electric field in the wave before reflection. The first subscript denotes whether the electric field in the incident wave is parallel ($||$) or perpendicular ($\perp$) to the plane of incidence and the second subscript refers in the same way to the reflected wave. A cartesian coordinate system (x, y, z) is now taken with z measured vertically upwards. The incident wave has its normal in the xz plane inclined at an angle θ to the z axis. The components of the electric field are $E_{||}$ in the xz plane and $E_{\perp}$ perpendicular to this plane (i.e., in the direction of increasing y). When the $+$ sign is taken in eq (1), the refractive index is denoted μ_0 corresponding to the ordinary wave, and when the $-$ sign is taken the refractive index is denoted μ_e corresponding to the extraordinary wave. With this convention, it can be shown that

$$E_{\perp 0}/E_{||0} = -i \text{ and } E_{\perp e}/E_{||e} = i \tag{10.2}$$

in the northern hemisphere.

The incident wave is now characterized by a factor $\exp[-ik(x \sin\theta + z\cos\theta)x]$ and the reflected wave, therefore, contains a factor $\exp[-ik(x \sin\theta - z\cos\theta)x]$. Furthermore, the transmitted waves have factors $\exp[-ik(x\sin\theta_0 + z\cos\theta_0)\mu_0 x]$ and $\exp[-ik(x\sin\theta_e + z\cos\theta_e)\mu_e x]$. The reflection coefficients are now obtained by matching tangential field components at the air-ionosphere interface. The results, expressed in a form suitable for computation are listed below.

$${}_{||}R_{||} = [(\mu_0+\mu_e)(C^2 - C_0C_e) + (\mu_0\mu_e - 1)(C_0 + C_e)C]/D \tag{10.3}$$

$${}_{||}R_{\perp} = 2iC(\mu_0 C_0 - \mu_e C_e)/D \tag{10.4}$$

$${}_{\perp}R_{||} = 2iC(\mu_0 C_e - \mu_e C_0)/D \tag{10.5}$$

$${}_{\perp}R_{\perp} = [(\mu_0+\mu_e)(C^2 - C_0C_e) - (\mu_0\mu_e - 1)(C_0 + C_e)C]/D \tag{10.6}$$

where

$$D=(\mu_0+\mu_e)(C^2+C_0C_e)+(\mu_0\mu_e+1)(C_0+C_e)C \tag{10.7}$$

and

$$C=\cos\theta, C_0=\cos\theta_0, C_e=\cos\theta_e.$$

Numerical values based on these formulas are available.

Now for highly oblique incidence the value of $|C|$ is small. Thus for $|C^2|\ll 1$

$${}_{||}R_{||}\cong-\left[\frac{1-\dfrac{(\mu_0\mu_e-1)(C_0+C_e)}{(\mu_0+\mu_e)C_0C_e}\,C}{1+\dfrac{(\mu_0\mu_e+1)(C_0+C_e)}{(\mu_0+\mu_e)C_0C_e}\,C}\right]. \tag{10.8}$$

Furthermore if $|\mu_0\mu_eC^2|\ll 1$,

$${}_{||}R_{||}\cong-\left[1-\frac{2\mu_0\mu_e}{\mu_0+\mu_e}\,\frac{C_0+C_e}{C_0C_e}\,C\right] \tag{10.9}$$

$$\cong-\exp(-2\beta_{||}C)$$

where

$$\beta_{||}=\frac{\mu_0\mu_e}{\mu_0+\mu_e}\,\frac{C_0+C_e}{C_0C_e}. \tag{10.10}$$

To the same approximation

$${}_{\perp}R_{\perp}\cong-\left[1-\frac{2}{\mu_0+\mu_e}\,\frac{C_0+C_e}{C_0C_e}\,C\right]\cong-\exp(-2\beta_{\perp}C) \tag{10.11}$$

where

$$\beta_{\perp}=\frac{1}{\mu_0+\mu_e}\,\frac{C_0+C_e}{C_0C_e}. \tag{10.12}$$

Also

$${}_{||}R_{\perp}\cong\frac{2iC}{C_0C_e}\,\frac{\mu_0C_0-\mu_eC_e}{\mu_0+\mu_e} \tag{10.13}$$

and

$${}_{\perp}R_{||}\cong\frac{2iC}{C_0C_e}\,\frac{\mu_0C_e-\mu_eC_0}{\mu_0+\mu_e}. \tag{10.14}$$

It is immediately evident from the above that as θ tends to $\pi/2$ (i.e., grazing incidence), the reflection coefficients ${}_{||}R_{||}$ and ${}_{\perp}R_{\perp}$ are both approaching -1 whereas the conversion coefficients ${}_{||}R_{\perp}$ and ${}_{\perp}R_{||}$ are both approaching zero. In this sense a sharply bounded ionosphere behaves as an equivalent isotropic medium for highly oblique incidence.

Some further simplifications are possible when the ionosphere is effectively a good conductor. For example, if

$$|\omega_r/\omega|\gg 1$$

$$\mu_0^2=1-i(\omega_r/\omega)e^{i\tau}\cong-i(\omega_r/\omega)e^{i\tau}, \tag{10.15}$$

$$\mu_e^2\cong 1-i(\omega_r/\omega)e^{-i\tau}\cong-i(\omega_r/\omega)e^{-i\tau}. \tag{10.16}$$

Consequently,

$${}^{||}_{\perp}R_{||}^{\perp}\cong\frac{(1+i)(\omega/\omega_r)^{\frac{1}{2}}\cos(\tau/2)(\mathrm{C}^2-1)\pm 2^{\frac{1}{2}}[1-i\omega/\omega_r]\,\mathrm{C}}{(1+i)(\omega/\omega_r)^{\frac{1}{2}}\cos(\tau/2)(\mathrm{C}^2+1)+2^{\frac{1}{2}}[1+i\omega/\omega_r]\,\mathrm{C}} \tag{10.17}$$

and

$${}^{||}_{||}R_{\perp}^{\perp}\cong\frac{-2(\omega/\omega_r)^{\frac{1}{2}}(1+i)\,\mathrm{C}\sin(\tau/2)}{(1+i)(\omega/\omega_r)^{\frac{1}{2}}\cos(\tau/2)(\mathrm{C}^2+1)+2^{\frac{1}{2}}[1+i\omega/\omega_r]\,\mathrm{C}}. \tag{10.18}$$

These may be further approximated, for $\cos\theta << 1$, by

$$1-{}_{\parallel}R_{\parallel} \cong 2\left(\frac{i\omega}{\omega_r}\right)^{\frac{1}{2}} \frac{\cos(\tau/2)}{\cos\theta},\quad {}_{\perp}R_{\perp}+1 \cong 2\left(\frac{i\omega}{\omega_r}\right)^{\frac{1}{2}} \frac{\cos\theta}{\cos(\tau/2)} \tag{10.19}$$

provided the right-hand sides of these equations are small compared to one. To the same approximation

$$\begin{matrix} {}_{\parallel}R_{\perp} \\ {}_{\perp}R_{\parallel} \end{matrix} \cong -2(i\omega/\omega_r)^{\frac{1}{2}} \sin(\tau/2). \tag{10.20}$$

The quasi-longitudinal approximation used above is only valid when

$$\frac{\omega_T^4}{4\omega^2\omega_L^2} << \left|\left(1-\frac{\omega_0^2}{\omega^2}-i\frac{\nu}{\omega}\right)^2\right|$$

where ω_L and ω_T are the longitudinal and transverse components of the (angular) gyro frequency. Clearly, this condition is violated when the transverse component of the earth's magnetic field is large such as for propagation around the magnetic equator.

The case of a purely horizontal and transverse field has been considered by N. F. Barber and D. D. Crombie (to be published in Journal of Atmospheric and Terrestrial Physics). Their results, applicable to a sharply bounded ionosphere, may be written in the following form

$${}_{\parallel}R_{\parallel} = \frac{C-\Delta}{C+\Delta},\quad {}_{\parallel}R_{\perp} = 0 \tag{10.21}$$

where

$$\Delta = \frac{\left[C^2+\dfrac{1+iL}{iL-L^2-\gamma^2}\right]^{\frac{1}{2}}(iL-L^2-\gamma^2)-i\gamma S}{(1+iL)^2-\gamma^2}$$

where $L=\omega/\omega_r$ and $\gamma=\omega_T\omega/\omega_0^2$. For east-to-west propagation (along the magnetic equator), γ is positive, while for west-to-east propagation, γ is negative. For highly oblique incidence, the above simplifies to

$$\frac{1}{\Delta} \cong \beta_{\parallel} \tag{10.22}$$

where

$$\frac{1}{\beta_{\parallel}} = \frac{(1+iL)^{\frac{1}{2}}(iL-L^2-\gamma^2)^{\frac{1}{2}}-i\gamma}{(1+iL)^2-\gamma^2}.$$

Furthermore if, $|\beta_{\parallel} C| << 1$

$${}_{\parallel}R_{\parallel} \cong -e^{-2\beta_{\parallel}C} \tag{10.23}$$

which has the same form as equation (11)

The exact determination of the reflection coefficients for any orientation of the earth's magnetic field may be carried out using a method outlined by Bremmer [3]. This has been done by Johler and Walters whose results are to be published in the following issue of this journal.

11. Mode Series for an Anisotropic Ionosphere

In this section the formalism for the mode theory is developed for a plane earth and an anisotropic sharply-bounded ionosphere. The geometry is the same as in section (3) where the

ionosphere was assumed to be isotropic. In the present case the reflection coefficient $[R_i]''$ is regarded as a matrix and written in the form

$$[R_i]''=\begin{bmatrix} {}_{\parallel}R_{\parallel} & {}_{\perp}R_{\parallel} \\ {}_{\parallel}R_{\perp} & {}_{\perp}R_{\perp} \end{bmatrix}'' \tag{11.1}$$

where the two primes are to indicate that it is a two column matrix. The individual coefficients ${}_{\parallel}R_{\parallel}$, ${}_{\perp}R_{\parallel}$, ${}_{\parallel}R_{\perp}$, and ${}_{\perp}R_{\perp}$ discussed earlier, indicate the complex ratio of an electric field component in the wave after reflection to an electric field component in the wave before reflection. The first subscript denotes whether the electric field specified in the incident wave is parallel ($\parallel$) or perpendicular ($\perp$) to the plane of incidence. The second subscript refers in the same way to the electric field in the reflected wave.

When the ionosphere becomes isotropic corresponding to a zero magnetic field, the reflection coefficient in matrix notation becomes simply

$$[R_i]''=\begin{bmatrix} R_i & 0 \\ 0 & R_i^h \end{bmatrix}'' \tag{11.2}$$

where R_i and R_i^h are the complex scalar reflection coefficients for vertically-polarized and horizontally-polarized waves, respectively. The corresponding [matrix] reflection for the ground is

$$[R_g]''=\begin{bmatrix} R_g & 0 \\ 0 & R_g^h \end{bmatrix}''. \tag{11.3}$$

The case of two successive reflections, the first from the ground and the second from the anisotropic ionosphere, is represented by the matrix

$$[R_gR_i]''=[R_i]''\times[R_g]''=\begin{bmatrix} R_g\,{}_{\parallel}R_{\parallel} & R_g^h\,{}_{\perp}R_{\parallel} \\ R_g\,{}_{\parallel}R_{\perp} & R_g^h\,{}_{\perp}R_{\perp} \end{bmatrix}'' \tag{11.4}$$

In the case of the isotropic ionosphere this reduces to

$$[R_gR_i]''=\begin{bmatrix} R_gR_i & 0 \\ 0 & R_g^hR_i^h \end{bmatrix}'' \tag{11.5}$$

The arguments employed here are virtually identical to those of Budden who, however, assumes a perfectly conducting ground, such that

$$R_g=+1 \text{ and } R_g^h=-1.$$

In the previous formulation for a vertical electric dipole between the plane ground surface and a sharply bounded isotropic ionosphere, the fields could be completely derived from an electric Hertz vector which has only a z component. Of course, if the source was not symmetrical it was necessary to introduce an additional component of the electric Hertz vector. When the upper boundary is anisotropic, the single component Hertz vector is not adequate even for a vertical electric dipole source. This is not surprising since the TM modes are coupled to the TE modes by the anisotropic boundary conditions.

Any electromagnetic field, in such a parallel plate region, can be obtained from a superposition of TM and TE modes which are derived from electric and magnetic Hertz vectors.

These are denoted individually by $\vec{\Pi}_\parallel$ and $\vec{\Pi}_\perp$ or, collectively by the matrix

$$[\vec{\Pi}]''=\begin{bmatrix}\vec{\Pi}_\parallel\\ \vec{\Pi}_\perp\end{bmatrix}' \tag{11.6}$$

where the single prime is to indicate that it is a single column matrix. $\vec{\Pi}_\perp$ which is a magnetic Hertz vector is often referred to as a Fitzgerald vector. Furthermore, the electric and magnetic fields can also be written as single column matrices in the manner

$$[\vec{E}]'=\begin{bmatrix}\vec{E}_\parallel\\ \vec{E}_\perp\end{bmatrix}',\ [\vec{H}]'=\begin{bmatrix}\vec{H}_\parallel\\ \vec{H}_\perp\end{bmatrix}' \tag{11.7}$$

where

$$\vec{E}_\parallel=(k^2+\text{grad div})\vec{\Pi}_\parallel \tag{11.8}$$

$$\vec{H}_\parallel=i\epsilon\omega\ \text{curl}\ \vec{\Pi}_\parallel \tag{11.9}$$

$$\eta\vec{H}_\perp=(k^2+\text{grad div})\vec{\Pi}_\perp \tag{11.10}$$

$$\eta\vec{E}_\perp=-i\mu\omega\ \text{curl}\ \vec{\Pi}_\perp. \tag{11.11}$$

The intrinsic impedance η is introduced in the latter two equations to make $\vec{\Pi}_\parallel$ and $\vec{\Pi}_\perp$ of the same dimensions. The Hertz vector in matrix form corresponding to the primary excitation is then written

$$[\vec{\Pi}_p]'=\begin{bmatrix}\vec{\Pi}_p\\ 0\end{bmatrix}' \tag{11.12}$$

where $\vec{\Pi}_p$ has only a z component $\Pi_z^{(p)}$. To match boundary conditions in the case of azimuthal symmetry it is only necessary that the vectors $\vec{\Pi}_\parallel$ and $\vec{\Pi}_\perp$ have a z component. The condition of azimuthal symmetry is achieved if the reflection coefficients themselves are independent of the azimuthal coordinate ϕ.

Formally the solution has the same form as the isotropic case if the appropriate reflection coefficients are now regarded as matrices. For example for the space $o<z<h$ the (matrix) Hertz vector

$$[\Pi_z]'=\frac{ik[M]'}{2}\int_\Gamma[F(C)]''H_0^{(2)}(kS\rho)dC \tag{11.13}$$

where

$$[M]=\begin{bmatrix}M\\ 0\end{bmatrix}'\ \text{with}\ M=\frac{I\,ds}{4\pi i\epsilon\omega}$$

and where

$$[F(C)]''=\frac{(e^{ikCz}+[R_g]''e^{-ikCz})(e^{ikC(h-z_0)}+[R_t]''e^{-ikC(h-z_0)})}{e^{ikCh}(1-[R_gR_t]''e^{-2ikhC})}. \tag{11.14}$$

It should be noted that the denominator in the above expression is also a two column matrix. Inverting this, following the usual rules for such operations, leads to

$$[F(C)]''=(e^{ikCz}+[R_g]''e^{-ikCz})(1/\Delta)[N]''\{e^{i2kCh}+[R_t]''\} \tag{11.15}$$

for $z_0=0$, where

$$\Delta=\begin{vmatrix}e^{2ikCh}-{}_\parallel R_\parallel\ R_g & {}_\perp R_\parallel\ \ R_g^h\\ {}_\parallel R_\perp\ R_g & e^{2ikCh}-{}_\perp R_\perp\ R_g^h\end{vmatrix} \tag{11.16}$$

and

$$[N]''=\begin{bmatrix} e^{2iCkh}-{}_{\perp}R_{\perp}\,R_g^h & {}_{\perp}R_{\parallel}\,R_g^h \\ {}_{\parallel}R_{\perp}\,R_g & e^{2iCkh}-{}_{\parallel}R_{\parallel}\,R_g \end{bmatrix}''. \tag{11.17}$$

The corresponding residue series are thus given by

$$\begin{bmatrix} \Pi_{\parallel} \\ \Pi_{\perp} \end{bmatrix}'=-\pi kM\sum_p \frac{e^{2iC_pkh}}{[\partial\Delta/\partial C]_{C=C_p}}H_0^{(2)}(kS_p\rho)\begin{bmatrix} G_{\parallel p} \\ G_{\perp p} \end{bmatrix}' \tag{11.18}$$

with

$$G_{\parallel p}=f_p(z)(e^{2iC_pkh}-R_g^h{}_{\perp}R_{\perp}) \tag{11.19}$$

and

$$G_{\perp p}=if_p^h(z)R_{g\parallel}R_{\perp}. \tag{11.20}$$

The summation is over the poles of the integrand $[F(C)]''$. Clearly this corresponds to the roots of the equation, $\Delta=0$, which are designated $C=C_p$. It is understood that all quantities in the summand of eq (18) are to be evaluated at $C=C_p$. The "height-factors" $f_p(z)$ and $f_p^h(z)$ have the usual form, that is

$$2f_p(z)=(R_g)^{-\frac{1}{2}}e^{ikC_pz}+(R_g)^{\frac{1}{2}}e^{-ikC_pz} \tag{11.21}$$

and

$$2f_p^h(z)=(R_g^h)^{-\frac{1}{2}}e^{ikC_pz}+(R_g^h)^{\frac{1}{2}}e^{-ikC_pz}. \tag{11.22}$$

The above results reduce to those of Budden when the ground is perfectly conducting.

The modes excited in the waveguide can be logically grouped into two sets. The first has a TM (transverse magnetic) character and the second has a TE (transverse electric) character. To obtain the attenuation and the phase constants of these individual modes, it is adequate to consider the anisotropy as sort of a perturbation to the corresponding TE and TM modes for the isotropic case.[6]

To simplify the discussion the ground is considered to be perfectly conducting. That is, $R_g=1$ and $R_g^h=-1$. The modal equation now becomes

$$(e^{i2kCh}-{}_{\parallel}R_{\parallel})(e^{i2kCh}+{}_{\perp}R_{\perp})+{}_{\parallel}R_{\perp}{}_{\perp}R_{\parallel}=0. \tag{11.23}$$

When mode coupling is disregarded this breaks into two equations

$${}_{\parallel}R_{\parallel}e^{-i2kCh}=1=e^{-i2\pi n} \tag{11.24a}$$

$${}_{\perp}R_{\perp}e^{-i2kCh}=-1=-e^{-i2\pi m} \tag{11.24b}$$

where n and m take integral values. As mentioned in the previous section the reflection coefficients for highly oblique incidence may be approximated by

$${}_{\parallel}R_{\parallel}\cong -e^{-2\beta_{\parallel}C} \tag{11.25}$$

and

$${}_{\perp}R_{\perp}\cong -e^{-2\beta}{}_{\perp}C \tag{11.26}$$

where to a first order, $\beta_{\parallel}$ and $\beta_{\perp}$ are independent of C. It thus follows that the first approximation (indicated by a superscript (1)) for the solutions of the modal equation are

$$C_p\cong C_n^{(1)}=\frac{\pi(n-\frac{1}{2})}{kh-i\beta_{\parallel}}\ (n=1, 2, \ldots) \tag{11.27}$$

for the TM modes, and

$$C_p\cong C_m^{(1)}=\frac{\pi m}{kh-i\beta_{\perp}}\ (m=1, 2, 3, \ldots) \tag{11.28}$$

for the TE modes. These have exactly the same form as when the ionosphere is assumed to be isotropic. The difference lies in the value of the coefficients $\beta_{\parallel}$ and $\beta_{\perp}$ which are functions of the earth's magnetic field.

[6] It should be noted that the negative order modes in the case of an anisotropic ionosphere are not the same as the positive order modes. That is, $C_{-1}\neq -C_0$, $C_{-2}\neq -C_1$, etc.

A second approximation to the mode equations is obtained in the following way. The modal equation is rewritten in the two equivalent forms

$$1-{}_{\|}R_{\|}e^{-i2kCh}=2\delta(C) \tag{11.29}$$

$$1+{}_{\perp}R_{\perp}e^{-i2kCh}=2\gamma(C) \tag{11.30}$$

where

$$2\delta(C)=-\frac{{}_{\|}R_{\perp}\,{}_{\perp}R_{\|}e^{-i2kCh}}{e^{i2kCh}+{}_{\perp}R_{\perp}} \tag{11.31}$$

and

$$2\gamma(C)=-\frac{{}_{\|}R_{\perp}\,{}_{\perp}R_{\|}e^{-i2kCh}}{e^{i2kCh}-{}_{\|}R_{\|}}. \tag{11.32}$$

It is to be expected that $\delta(C)$ is a small quantity for the TM type modes and $\gamma(C)$ is a small quantity for the TE type modes. The second approximations then are obtained by replacing $\delta(C)$ by $\delta(C_n^{(1)})$ for the TM set, and replacing $\gamma(C)$ by $\gamma(C_m^{(1)})$ for the TE set. Solving eqs (29) and (30) with these substitutions leads readily to

$$C_p\cong C_n^{(2)}=\frac{\pi(n-\frac{1}{2})-i\delta(C_n^{(1)})}{kh-i\beta_{\|}} \tag{11.33}$$

for the TM type modes, and

$$C_p\cong C_m^{(2)}=\frac{\pi m-i\gamma(C_m^{(1)})}{kh-i\beta_{\perp}}. \tag{11.34}$$

for the TE type modes.

These should be adequate solutions since $|\delta(C_n{}^{(1)})|$ and $|\gamma(C_m{}^{(1)})|$ are small compared to unity for the important modes. In fact for most cases of practical interest, the first-order approximations should suffice.

To provide some idea of the character of the TM and the TE type modes excited by a vertical dipole source, the ratio of the tangential magnetic field in the two principal planes is considered. For the pth mode, this ratio is given by

$$\left(\frac{H_\rho}{H_\phi}\right)=\left(\frac{\partial^2\Pi_{\perp p}}{\partial\rho\partial z}\right)\Big/\left(-i\varepsilon\omega\,\frac{\partial\Pi_{\|p}}{\partial\rho}\right) \tag{11.35}$$

$$\cong=\frac{(\partial\Pi_{\perp p}/\partial z)}{ik\Pi_{\|p}}\text{ for }k\rho>>1$$

$$\cong-\frac{C_{p\|}R_{\perp}R_g}{[e^{2iCkh}-R_{g\perp}^{h}R_{\perp}]}\left\{\frac{\dfrac{1}{kC_p}\dfrac{\partial}{\partial z}f_p^h(z)}{f_p(z)}\right\} \tag{11.36}$$

where it is understood that the reflection coefficients are to be evaluated at $C=C_p$. The preceding expression reduces to

$$\left(\frac{H_\rho}{H_\phi}\right)\cong\frac{C_{p\|}R_{\perp}}{e^{2iCkh}+{}_{\perp}R_{\perp}} \tag{11.37}$$

which was given originally by Budden [11]. In general this ratio is small except near the TE resonance wherein the denominator becomes very small. For the important TM modes which are of low order, both C_p and ${}_{\|}R_{\perp}$ are small compared to unity and thus the magnetic field H_ρ in the direction of propagation has a relatively small magnitude.

12. Higher Approximations to the Curved Earth Theory

In the previous sections the mode series for a concentric spherical earth-air-ionosphere system was developed. In order to simplify the discussion and lead to results suitable for

immediate use, rather crude approximations were introduced. In this section the problem is reformulated in a more rigorous fashion and higher order approximations for the various spherical wave functions are introduced. This analysis is really an extension of the work of Watson, Rydbeck, and Bremmer. The final results indicate the range of validity of the lower order approximations used in the earlier sections. The formulas are in a form which is suitable for numerical computation.

The earth is represented by a homogeneous sphere of radius a and is surrounded by a concentric homogeneous sharply-bounded ionosphere of radius c. The source is a vertical electric dipole of strength $I\,ds$ and is located at $r=b$. The electrical constants of the air space are denoted ϵ and μ and subscripts g and i are added to these when reference is made to the ground and the ionosphere, respectively (see fig. 5).

The fields can be expressed in terms of a Hertz vector which has only a radial component U, and thus, for the region $a<r<a+h$.

$$E_r=\left(k^2+\frac{\partial^2}{\partial r^2}\right)(rU) \qquad H_r=0$$

$$E_\theta=\frac{1}{r}\frac{\partial^2}{\partial r\partial\theta}(rU) \qquad H_\theta=0$$

$$E_\phi=0 \qquad H_\phi=-i\epsilon\omega\frac{\partial U}{\partial\theta} \tag{12.1}$$

where $k=(\epsilon\mu)^{1/2}\omega$.[7] A subscript g and i are also added to the field quantities when reference is made to the regions $r<a$ and $r>c$, respectively. Furthermore, $k_g=(\epsilon_g\mu_g)^{1/2}\omega$ and $k_i=(\epsilon_i\mu_i)^{1/2}\omega$ are the respective wave numbers for these two regions.

The Hertz functions satisfy the inhomogeneous wave equation

$$(\nabla^2+k^2)\,U=\overline{C}\frac{\delta(r-b)\,\delta(\theta)}{2\pi r^2\sin\theta} \tag{12.2}$$

for $a<r<a+h$, where the δ's are unit impulse functions. The factor $2\pi r^2\sin\theta$ is the Jacobian of the transformation from rectangular to spherical coordinates. The constant $\overline{C}$ is to be chosen so that U has the proper singularity at the dipole, that is

$$-bU\rightarrow\frac{e^{-ikR}}{4\pi i\omega\epsilon R}I\,ds \text{ for } R\rightarrow 0$$

where $R=[r^2+b^2-2br\cos\theta]^{1/2}$, and therefore $\overline{C}=(i/\omega\epsilon)I\,ds$.

The field in the region $a<r<a+h$ is now written as the sum of the two parts U_e+U_s, where U_e has the proper dipole singularity at $R=0$, and U_s is finite at the point. As U_s is a solution of the homogeneous wave equation, it can be written in the form

$$U_s=\frac{ik\overline{C}}{4\pi}\sum_{q=0}^{\infty}(2q+1)\,[A_q h_q^{(2)}(kr)+B_q j_q(kr)]\,P_q(\cos\theta) \tag{12.3}$$

where $j_q(kr)$ and $h_q^{(2)}(kr)$ are spherical Hankel functions of the first and fourth kind, respectively, and $P_q(\cos\theta)$ is the Legendre function. The summation is over positive integral values of q. The corresponding expression for U_e is given by

$$U_e=\frac{ik\overline{C}}{4\pi}\sum_{q=0}^{\infty}(2q+1)\,P_q(\cos\theta)\begin{cases} j_q(kr)\,h_q^{(2)}(kb); & \text{for } r<b \\ h_q^{(2)}(kr)\,j_q(kb); & \text{for } r>b. \end{cases} \tag{12.4}$$

[7] The function $U=-i\eta\psi$ in terms of scalar function ψ used previously for the potential.

Since there are no singularities other than the source dipole, the Hertz functions U_g and U_i are solutions of the homogeneous wave equations

$$(\nabla^2+k_g^2)(rU_g)=0 \qquad \text{for } 0\leqq r\leqq a \tag{12.5}$$

and

$$(\nabla^2+k_i^2)(rU_i)=0 \qquad \text{for } r\geqq c. \tag{12.6}$$

Noting that U_g is to be finite at $r=0$ the solution must be of the form

$$U_g=\frac{ik\bar{C}}{4\pi}\sum_{q=0}^{\infty}(2q+1)P_q(\cos\theta)a_q j_q(k_g r) \tag{12.7}$$

where a_q is a coefficient which is independent of r and θ. Furthermore, since U_i is to give rise to an outgoing wave at $r=\infty$, the solution is of the form

$$U_i=\frac{ik\bar{C}}{4\pi}\sum_{q=0}^{\infty}(2q+1)P_q(\cos\theta)b_q h_q^{(2)}(k_i r) \tag{12.8}$$

where b_q is a coefficient.

The four unknown coefficients A_q, B_q, a_q, and b_q, can be found from the boundary conditions at $r=a$ and c. These require the continuity of the tangential field components. In order to facilitate the solution and to readily permit later generalizations, the (four) boundary conditions as stated above can be replaced by two impedance type boundary conditions. For the qth terms of the series these read

$$E_\theta^{(q)}=-Z_g^{(q)}H_\phi^{(q)} \quad \text{at} \quad r=a \tag{12.9}$$

and

$$E_\theta^{(q)}=Z_i^{(q)}H_\phi^{(q)} \quad \text{at} \quad r=c \tag{12.10}$$

where

$$Z_g^{(q)}=\frac{1}{i\epsilon\omega}\frac{\frac{\partial}{\partial r}[rj_q(k_g r)]}{rj_q(k_g r)}$$

and

$$Z_i^{(q)}=-\frac{1}{i\epsilon\omega}\frac{\frac{\partial}{\partial r}[rh_q^{(2)}(k_i r)]}{rh_q^{(2)}(k_i r)}.$$

Replacing kr by x, eqs (9) and (10) may be rewritten

$$\frac{1}{x}\frac{\partial}{\partial x}(xU)=i(Z_g^{(q)}/\eta)U \qquad \text{for } x=ka \tag{12.11}$$

and

$$\frac{1}{x}\frac{\partial}{\partial x}(xU)=-i(Z_i^{(q)}/\eta)U \qquad \text{for } x=kc. \tag{12.12}$$

Applying these to eq (3) enable A_q and B_q to be obtained explicitly in terms of known quantities. Using these results leads readily to the following exact solution for $a\leqq r\leqq b$.

$$U_s=\frac{ik\bar{C}}{4\pi}\sum_{q=0}^{\infty}(2q+1)h_q^{(2)}(kb)h_q^{(1)}(kr)P_q(\cos\theta)\frac{F_q}{D_q} \tag{12.13}$$

where

$$F_q=\left[1+R_g^{(q)}\frac{h_q^{(1)}(ka)}{h_q^{(2)}(ka)}\frac{h_q^{(2)}(kr)}{h_q^{(1)}(kr)}\right]\left[1+R_i^{(q)}\frac{h_q^{(2)}(kc)}{h_q^{(1)}(kc)}\frac{h_q^{(1)}(kb)}{h_q^{(2)}(kb)}\right], \tag{12.14}$$

$$D_q=1-R_g^{(q)}R_i^{(q)}\frac{h_q^{(1)}(ka)}{h_q^{(2)}(ka)}\frac{h_q^{(2)}(kc)}{h_q^{(1)}(kc)} \tag{12.15}$$

$$R_g^{(q)}=-\frac{ln'[ka\,h_q^{(1)}(ka)]-iZ_g^{(q)}/\eta}{ln'[ka\,h_q^{(2)}(ka)]-iZ_g^{(q)}/\eta} \tag{12.16}$$

$$R_i^{(q)}=-\frac{ln'[kc\,h_q^{(2)}(kc)]+iZ_i^{(q)}/\eta}{ln'[kc\,h_q^{(1)}(kc)]+iZ_i^{(q)}/\eta}. \tag{12.17}$$

The symbol ln' denotes logarithmic differentiation, for example

$$ln'[ka\,h_q^{(1)}(ka)]=\left.\frac{\frac{\partial}{\partial x}[xh_q^{(1)}(x)]}{xh_q^{(1)}(x)}\right]_{x=ka}. \tag{12.18}$$

The above result, although rigorous, is not of practical value for vlf propagation calculations because of poor convergence of the series solutions. In fact, something of the order of 2 ka terms are required to achieve 5 percent accuracy. At 15 kc, for example, 2 $ka\cong 2{,}000$ which is a rather large number. An important observation, however, is that terms of order q beyond 2 ka contribute little to the series. Thus the spherical Bessel functions $j_q(k_ga)$ may be replaced by the Debye or second-order approximation since $|k_ga|>>q$ in the important range of q so long as $|k_g|>>k$ (i.e., well conducting ground). Thus

$$ln'[k_gaj_q(k_ga)]\cong i\left[1-\frac{q^2}{k_g^2}\right]^{\frac{1}{2}}. \tag{12.19}$$

Similarly, for $|k_i|>>k$

$$ln'[k_iah_q^{(2)}(k_ia)]\cong -i\left[1-\frac{q^2}{k_i^2}\right]^{\frac{1}{2}}. \tag{12.20}$$

Since the total field is of the form

$$U=\sum_{q=0}^{\infty}(2q+1)f(q)P_q(\cos\theta) \tag{12.21}$$

it can be rewritten as a contour integral over q where the integrand has poles when q takes integrand values. Such a representation is

$$U=i\int_{C_1+C_2}\frac{qdq}{\cos q\pi}f(q-\tfrac{1}{2})P_{q-\frac{1}{2}}[\cos(\pi-\theta)] \tag{12.22}$$

where the contour C_1+C_2 encloses the real axis as illustrated in figure 9. Noting that the poles of the integrand are located at $q=1/2, 3/2, 5/2$ . . . etc., it can be readily verified by the theorem of residues that this integral is equivalent to eq (21). Now, subject to the validity of the second-order approximations, for the wave functions of order k_ga and k_ia mentioned above, the function $f(q-\frac{1}{2})$ is an even function of q. This means that the part of the contour C_1 just above the positive real axis can be replaced by C_1' which is located just below the negative real axis (see fig. 9). The contour $C_1'+C_2$ is now entirely equivalent to L, a straight line running along just below the real axis. Replacing $q-\frac{1}{2}$ by ν the contour representation for U takes the form

$$U=-i\int_L\frac{(\nu+\frac{1}{2})}{\sin\nu\pi}f(\nu)P_\nu[\cos(\pi-\theta)]\,d\nu. \tag{12.23}$$

It is to be noted that this manipulation of the contours is only strictly justified when $f(q-\frac{1}{2})$ is an even function of q. This is well justified when $|k_2|^2$ and $|k_1|^2$ are both $>>k^2$.

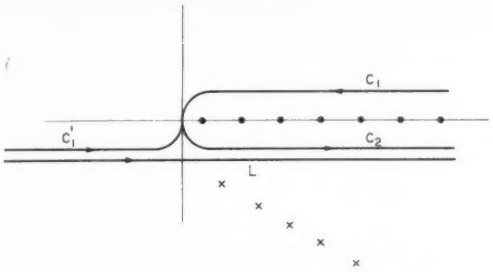


FIGURE 9. *The contours in the complex* q *plane.*

The next step in the analysis is to close L by an infinite semicircle in the negative half-plane. The contribution from this part of the contour vanishes as the radius of the semicircle approaches infinity because of the exponentially decreasing character of the integrand. The value of the integral for U along the contour L is now equal to $-2\pi i$ times the sum of the residues of the integrand evaluated at the poles of $f(\nu)$ located in the lower half-plane of ν. It then follows that

$$U=-ik\bar{C}\sum_{\nu}\frac{(\nu+\frac{1}{2})}{\sin\nu\pi}h_{\nu}^{(2)}(kb)h_{\nu}^{(1)}(kr)\frac{F_{\nu}}{D_{\nu}'}P_{\nu}[\cos(\pi-\theta)] \tag{12.24}$$

where $D_{\nu}^{(\prime)}=\partial D_{\nu}/\partial\nu$. All quantities in the summand are to be evaluated at the poles of $f(\nu)$ which are the roots of the equation

$$D_{\nu}=0.$$

This equation is precisely the same as the one discussed earlier (*i.e.*, eq (5.5)). At that time the relevant spherical wave functions of order ka and kc were simplified by the use of the Debye or second order approximation. The Hankel or third order approximation will now be employed. It may be written [26]

$$xh_{\nu}^{(1)}_{(2)}(x)\cong\left(\frac{\pi r}{6}\right)^{\frac{1}{2}}\frac{x}{\nu+\frac{1}{2}}\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{\frac{3}{4}}H_{1/3}^{(2)}(\rho)\exp\left[\pm i\left(\frac{5}{12}\pi-\rho\right)\right]S_{\nu}^{(1)}_{(2)}(x) \tag{12.25}$$

where

$$S_{\nu}^{(1)}_{(2)}(x)=\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{-\frac{1}{4}}\exp\left(\mp i\frac{\pi}{4}\right)\exp\left[\pm i\int_{\nu+\frac{1}{2}}^{x}\left(1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right)^{\frac{1}{2}}dx\right] \tag{12.26}$$

and $H_{1/3}^{(1)}_{(2)}(\rho)$ is the Hankel function of order 1/3 of argument ρ given by

$$\rho=\frac{\nu+\frac{1}{2}}{3}\left[\frac{x^2}{(\nu+\frac{1}{2})^2}-1\right]^{\frac{3}{2}}. \tag{12.27}$$

For Re $(\nu+\frac{1}{2})<<x$ or for $|\rho|>>1$, the above reduces to

$$xh_{\nu}^{(1)}_{(2)}(x)\cong S_{\nu}^{(1)}_{(2)}(x) \tag{12.28}$$

which is the Debye approximation used in section (5).

The third order representation for the logarithmic derivative is

$$ln'[xh_{\nu}^{(1)}_{(2)}(x)]\cong e^{\pm i\frac{2\pi}{3}}\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{\frac{1}{2}}\frac{H_{2/3}^{(1)}_{(2)}(\rho)}{H_{1/3}^{(1)}_{(2)}(\rho)} \tag{12.29}$$

while the corresponding second order approximation, valid for $|\rho|>>1$ is simply

$$ln'[xh_{\nu}^{(1)}_{(2)}(x)]\cong\pm i\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{\frac{1}{2}}. \tag{12.30}$$

For convenience in what follows, it is desirable to introduce two new spherical reflection coefficients r_g and r_i which are connected to R_g and R_i in the following manner:

$$r_g=-\frac{ln'[kah_\nu^{(2)}(ka)]}{ln'[kah_\nu^{(1)}(ka)]}\,R_g=\frac{1-\dfrac{i\Delta_g}{ln'[kah_\nu^{(1)}(ka)]}}{1-\dfrac{i\Delta_g}{ln'[kah_\nu^{(2)}(ka)]}} \tag{12.31}$$

and

$$r_i=-\frac{ln'[kch_\nu^{(1)}(kc)]}{ln'[kch_\nu^{(2)}(kc)]}\,R_i=\frac{1+\dfrac{i\Delta_i}{ln'[kch_\nu^{(2)}(kc)]}}{1+\dfrac{i\Delta_i}{ln'[kch_\nu^{(1)}(kc)]}} \tag{12.32}$$

where

$$\Delta_g\cong\frac{k}{k_g}\left[1-\left(\frac{\nu+\frac{1}{2}}{k_ga}\right)^2\right]^{\frac{1}{2}} \tag{12.33}$$

and

$$\Delta_i\cong\frac{k}{k_i}\left[1-\left(\frac{\nu+\frac{1}{2}}{k_ia}\right)^2\right]^{\frac{1}{2}}. \tag{12.34}$$

These new reflection coefficients may be expressed to a high order of approximation by using the Hankel or third-order approximation for the spherical Bessel functions of argument ka or kc. Thus

$$r_g\cong\frac{1-\dfrac{\Delta_g}{\left[1-\left(\frac{\nu+\frac{1}{2}}{ka}\right)^2\right]^{\frac{1}{2}}}\,e^{-i\frac{\pi}{6}}\,\dfrac{H_{1/3}^{(1)}(\rho_a)}{H_{2/3}^{(1)}(\rho_a)}}{1+\dfrac{\Delta_g}{\left[1-\left(\frac{\nu+\frac{1}{2}}{ka}\right)^2\right]^{\frac{1}{2}}}\,e^{i\frac{\pi}{6}}\,\dfrac{H_{1/3}^{(2)}(\rho_a)}{H_{2/3}^{(2)}(\rho_a)}} \tag{12.35}$$

and

$$r_i\cong\frac{1-\dfrac{\Delta_i}{\left[1-\left(\frac{\nu+\frac{1}{2}}{kc}\right)^2\right]^{\frac{1}{2}}}\,e^{i\frac{\pi}{6}}\,\dfrac{H_{1/3}^{(2)}(\rho_c)}{H_{2/3}^{(2)}(\rho_c)}}{1+\dfrac{\Delta_i}{\left[1-\left(\frac{\nu+\frac{1}{2}}{kc}\right)^2\right]^{\frac{1}{2}}}\,e^{-i\frac{\pi}{6}}\,\dfrac{H_{1/3}^{(1)}(\rho_c)}{H_{2/3}^{(1)}(\rho_c)}} \tag{12.36}$$

where

$$\rho_a=\frac{\nu+\frac{1}{2}}{3}\left[\frac{(ka)^2}{(\nu+\frac{1}{2})^2}-1\right]^{3/2} \tag{12.37}$$

and

$$\rho_c=\frac{\nu+\frac{1}{2}}{3}\left[\frac{(kc)^2}{(\nu+\frac{1}{2})^2}-1\right]^{3/2}. \tag{12.38}$$

To this same approximation

$$r_g\cong e^{-i4\pi/3}\,\frac{H_{2/3}^{(2)}(\rho_a)}{H_{1/3}^{(2)}(\rho_a)}\,\frac{H_{1/3}^{(1)}(\rho_a)}{H_{2/3}^{(1)}(\rho_a)}\,R_g \tag{12.39}$$

and

$$r_i\cong e^{i4\pi/3}\,\frac{H_{2/3}^{(1)}(\rho_c)}{H_{1/3}^{(1)}(\rho_c)}\,\frac{H_{1/3}^{(2)}(\rho_c)}{H_{2/3}^{(2)}(\rho_c)}\,R_i. \tag{12.40}$$

When $|\rho_a|$ and $|\rho_c|>>1$

$$r_g\cong R_g\cong\frac{\left[1-\left(\frac{\nu+\frac{1}{2}}{ka}\right)^2\right]^{\frac{1}{2}}-\Delta_g}{\left[1-\left(\frac{\nu+\frac{1}{2}}{ka}\right)^2\right]^{\frac{1}{2}}+\Delta_g} \tag{12.41}$$

and

$$r_t \cong R_t \cong \frac{\left[1-\left(\frac{\nu+\frac{1}{2}}{kc}\right)^2\right]^{\frac{1}{2}}-\Delta_t}{\left[1-\left(\frac{\nu+\frac{1}{2}}{kc}\right)^2\right]^{\frac{1}{2}}+\Delta_t}. \tag{12.42}$$

On writing $\nu+\frac{1}{2}=kaS=kcS'$, these latter forms are readily identified as the Fresnel reflection coefficients

$$r_g \cong R_g \cong \frac{N_g^2(1-S^2)^{\frac{1}{2}}-(N_g^2-S^2)^{\frac{1}{2}}}{N_g^2(1-S^2)^{\frac{1}{2}}+(N_g^2-S^2)^{\frac{1}{2}}} \tag{12.43}$$

and

$$r_t \cong R_t \cong \frac{N_t^2[1-(S')^2]^{\frac{1}{2}}-[N_t^2-(S')^2]^{\frac{1}{2}}}{N_t^2[1-(S')^2]^{\frac{1}{2}}+[N_t^2-(S')^2]^{\frac{1}{2}}}. \tag{12.44}$$

Attention is turned specifically to the determination of the roots of the equation

$$D_\nu=0. \tag{12.45}$$

This may be written

$$R_g R_t \frac{h_\nu^{(1)}(ka)h_\nu^{(2)}(kc)}{h_\nu^{(2)}(ka)h_\nu^{(1)}(kc)}=e^{-2\pi i n} \tag{12.46}$$

where R_g and R_t are defined by eq (31) and (32) and n may take integral values. Emplying the third-order approximations, this may be rewritten, for Re $(\nu+\frac{1}{2})<ka$

$$r_g r_t \frac{H_{2/3}^{(2)}(\rho_c)H_{2/3}^{(2)}(\rho_a)}{H_{2/3}^{(1)}(\rho_c)H_{2/3}^{(2)}(\rho_a)} e^{-i2(\gamma_c-\gamma_a)}e^{i2(\rho_c-\rho_a)}=e^{-i2\pi n} \tag{12.47}$$

where

$$\gamma_a=\int_{\nu+\frac{1}{2}}^{ka}\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{\frac{1}{2}}dx \tag{12.48}$$

and

$$\gamma_c=\int_{\nu+\frac{1}{2}}^{kc}\left[1-\frac{(\nu+\frac{1}{2})^2}{x^2}\right]^{\frac{1}{2}}dx \tag{12.49}$$

while, for $kc>\text{Re }(\nu+\frac{1}{2})>ka$,

$$r_g r_t \frac{H_{2/3}^{(2)}(\rho_c)H_{2/3}^{(2)}(\rho_a)}{H_{2/3}^{(1)}(\rho_c)H_{2/3}^{(1)}(\rho_a)} e^{-i2\gamma_c}e^{i2\rho_c}=e^{-i2\pi n}. \tag{12.50}$$

In the above formulas the (spherical) reflection coefficients r_g and r_t are defined by eqs (35) and (36). When $|\rho_a|$ and $|\rho_c| >>1$ or if Re $(\nu+\frac{1}{2})<<ka$, the relevant equation for the modes is simply

$$r_g r_t e^{-i2(\gamma_c-\gamma_a)}=e^{-i2\pi n} \tag{12.51}$$

where r_g and r_t are defined by eq (41) and (42) which are the Fresnel form of the reflection coefficients. Equation (51) is identical to eq (5.10) which was discussed previously.

13. Influence of Stratification at the Lower Edge of the Ionosphere

Attention in previous sections has been largely confined to a sharply bounded homogeneous ionosphere. In view of the general uncertainty about the electrical properties of the lower edge of the ionosphere, a more elaborate model might hardly seem worthwhile. Furthermore, despite the geometrical simplicity of the above models, the computation of the modes is very

involved in the general case. Despite these disparaging remarks the inhomogeneity of the lower ionosphere may be considered in some cases without greatly increasing the complexity. Some of these generalizations are discussed here for what they are worth.

The theoretical treatment given in section (3) for a vertical electric dipole located in the air space between a flat ground and a plane interface of a homogeneous ionosphere may be easily generalized to a stratified ionosphere. The essential modification is to replace the ionosphere reflection coefficient $R_i(C)$ by a more elaborate form which is denoted $\overline{R}_i(C)$. For example, a two layer ionosphere is chosen. The lower edge is at $z=h$ and from there to $z=h+s$, the refractive index (assumed constant and isotropic) is N_1; at this point the refractive index (also assumed constant and isotropic) is N_2 and remains at this value thereafter. It is not at all difficult to show that, for vertical polarization [31],

$$\overline{R}_i(C)=\frac{N_1^2C-(N_1^2-S^2)^{\frac{1}{2}}Q}{N_1^2C+(N_1^2-S^2)^{\frac{1}{2}}Q} \tag{13.1}$$

where

$$Q=\frac{N_1^2(N_2^2-S^2)^{\frac{1}{2}}+N_2^2(N_1^2-S^2)^{\frac{1}{2}}\tanh\,[iks(N_1^2-S^2)^{\frac{1}{2}}]}{N_2^2(N_1^2-S^2)^{\frac{1}{2}}+N_1^2(N_2^2-S^2)^{\frac{1}{2}}\tanh\,[iks(N_1^2-S^2)^{\frac{1}{2}}]}. \tag{13.2}$$

Here it may be observed that if $|kN_1s| \ll 1$ the reflection coefficient becomes

$$\overline{R}_i(C)\cong\frac{N_1^2C-(N_2^2-S^2)^{\frac{1}{2}}}{N_2^2C+(N_2^2-S^2)^{\frac{1}{2}}}. \tag{13.3}$$

whereas if $|kN_1s| \gg 1$, it becomes

$$\overline{R}_i(C)\cong\frac{N_1^2C-(N_1^2-S^2)^{\frac{1}{2}}}{N_1^2C+(N_1^2-S^2)^{\frac{1}{2}}}. \tag{13.4}$$

These two limiting cases correspond respectively, to the conditions of an electrically thin and an electrically thick stratum. In the former case the effective reflection level is at $z=h+s$ and in the latter case it is at $z=h$.

The formula for $\overline{R}_i(C)$ may easily be generalized to any number of layers. For example, in the case of discrete layers or strata, $0<z<h$ corresponds to the air; $h<z<h+s_1$ corresponds to a stratum with index N_1, $h+s_1<z<h+s_1+s_2$ corresponds to a stratum with index N_2, and so on. (See fig. 10.) With this generalization, Q, in eq (1), is to be replaced by

$$Q_1=\frac{N_1^2(N_2^2-S^2)^{\frac{1}{2}}Q_2+N_2^2(N_1^2-S^2)^{\frac{1}{2}}\tanh\,[iks_1(N_1^2-S^2)^{\frac{1}{2}}]}{N_2^2(N_1^2-S^2)^{\frac{1}{2}}+N_1^2(N_2^2-S^2)^{\frac{1}{2}}Q_2\tanh\,[iks_1(N_1^2-S^2)^{\frac{1}{2}}]} \tag{13.5}$$

$$Q_2=\frac{N_2^2(N_3^2-S^2)Q_3+N_3^2(N_2^2-S^2)\tanh\,[iks_2(N_2^2-S^2)^{\frac{1}{2}}]}{N_3^2(N_2^2-S^2)+N_2^2(N_3^2-S^2)Q_3\tanh\,[iks_2(N_2^2-S^2)^{\frac{1}{2}}]} \tag{13.6}$$

and so on. Q_3, Q_4, Q_5 . . . are obtained by cyclic permutation of indices. It should be noted, however, for M discrete strata that $Q_M=1$ since effectively $s_M=\infty$. The resultant Hertz vector for the air space $0<z<h$ is then formally given by eq (3.13) with the more general meaning now attached to the ionosphere reflection coefficient. In the general case, the rigorous evaluation of the integrals would be extremely involved. However, using arguments similar to those for the homogeneous ionosphere, the field may be approximated as a sum of residues evaluated at the poles of the integrand. Thus the contributions from the branch points are again neglected since for finitely conducting layers they correspond to heavily damped waves. Therefore, the residue series formula given by eq (3.15) is also applicable if the reflection co-

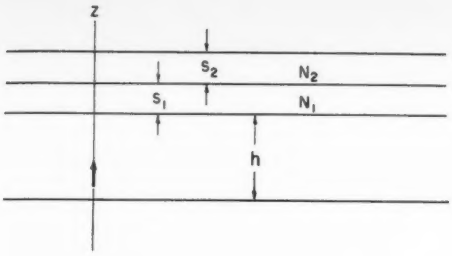


FIGURE 10. *Stratified model for ionosphere.*

efficient $R_t(C_n)$ is replaced by $\overline{R}_t(C_n)$. The modal equation now reads

$$\overline{R}_t(C_n)R_g(C_n)\exp(-2ikhC_n)=\exp(-i2\pi n) \tag{13.7}$$

where n takes integral values.

A numerical treatment and an application for the special case of a two-layer ionosphere has been carried out and reported in the literature [21]. There is no intrinsic difficulty in extending such calculations to an indefinitely large number of such layers each with infinitesimal thickness. For finitely conducting strata such a process converges and leads to an adequate representation for a continuous refractive index profile.

A great simplification to the formulas for a stratified ionosphere is effected if the refractive indices for all layers are large. For example, if $|N_1|, |N_2| \ldots |N_M| >> 1$, then

$$Q_1 \cong \frac{N_1Q_2+N_2\tanh(iks_1N_1)}{N_2+N_1Q_2\tanh(iks_2N_1)} \tag{13.8}$$

$$Q_2 \cong \frac{N_2Q_3+N_3\tanh(iks_2N_2)}{N_3+N_2Q_3\tanh(iks_2N_2)} \tag{13.9}$$

and so on. Thus to this approximation, Q_1 does not depend on the angle of incidence or the factor C. In this case, the modal equation simplifies to

$$khC_n \cong \pi n + i\Delta/C_n \tag{13.10}$$

where

$$\Delta=\frac{1}{N_g}+\frac{Q_1}{N_1}. \tag{13.11}$$

Regarding Δ/C_n as a small quantity, this can be solved to give

$$S_n=\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{\frac{1}{2}}-i\frac{\epsilon_n}{2kh}\Delta\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{-\frac{1}{2}} \tag{13.12}$$

where $\epsilon_0=1$, $\epsilon_n=2(n\neq 0)$. This is valid if $|\Delta kh| << 1$ and $|\Delta| << kh\,[1-(\pi n/kh)^2]$. Thus, at extremely low frequencies and for highly conducting layers, the propagation factor S_n is expressible in a relatively simple form.

The special case of a two layer ionosphere was considered in some detail in a previous paper [21]. Such a model was sufficient to explain the variation of the observed attenuation rate for frequencies of the order of 500 cps to 15 kc.

Exponential Profiles

At the extremely low frequencies it was seen that for a stratified model of the ionosphere the factor Q does not depend on C or S. In fact, it is not difficult to show that the surface impedance Z at $z=h$ looking outwards is given by

$$Z=\eta_0Q_1/N_1 \qquad \text{where } \eta_0 \cong 120\pi.$$

From this viewpoint is becomes quite easy to write down expressions for Q_1 for other profiles. For example, if the refractive index varies in an exponential fashion relatively simple formulas are obtained. Two cases which could be considered are

$$N(z)=1.0 \text{ for } 0<z<h,$$

$$N(z)=\overline{N} \exp\left[\mp(z-h)/l\right] \text{ for } z>h,$$

where the $(-)$ sign corresponds to a refractive index decreasing with height and the $(+)$ sign corresponds to a refractive index increasing with height. In the above, l is a scale factor; for example, at within a distance l above the lower edge of the ionosphere, the refractive index has changed from $\overline{N}$ to $\overline{N}/e$ or $\overline{N}e$ for the two respective cases.

Within the layers of the ionosphere propagation is vertically upwards and thus a component E of the electric field satisfies

$$\frac{\partial^2 E}{\partial z^2}+k^2N^2(z)E=0 \qquad \text{for } z>h.$$

Solutions of this equation for the exponential form of $N^2(z)$ are

$$E=\begin{matrix}\text{const}\times I_0\ (ik\overline{N}le^{-(z-h)/l}) \\ \text{const}\times K_0(ik\overline{N}le^{+(z-h)/l})\end{matrix} \tag{13.13}$$

for the two respective cases, where I_0 and K_0 are modified Bessel functions. The transverse component of the magnetic field is then found from Maxwell's equations, e.g., $i\mu\omega H=\partial E/\partial z$. The surface impedance Z at the lower edge of this model of the ionosphere is then defined by

$$Z=E/H]_{z=h}.$$

For the case when the refractive index *decreases* with height

$$Z=\frac{\eta_0}{N_1}Q, \text{ with } Q=\frac{I_0(ik\overline{N}l)}{I_1(ik\overline{N}l)} \tag{13.14}$$

and when the refractive index increases with height,

$$Z=\frac{\eta_0}{N_1}Q, \text{ with } Q=\frac{K_0(ik\overline{N}l)}{K_1(ik\overline{N}l)}. \tag{13.15}$$

For low frequencies, $ik\overline{N}l\cong\sqrt{i}\,x$ with $x=|\overline{N}|\,kl$. The arguments of modified Bessel functions are thus proportional to $\sqrt{i}$. Numerical values are shown in the table for certain real values of x. In both these cases, it may be observed that as the scale length l approaches infinity, Z becomes $\eta_0/\overline{N}$ as it should.

The simplicity of the above formulas for the exponential profiles is due to the inherent assumption that the refractive index $\overline{N}(z)$ is large compared to unity for $z>h$. When this condition is violated, such as it would be at frequencies above several kilocycles per second, the solution is not expressible in closed form, except for horizontal polarization [32] which is really only of academic interest at vlf. Nevertheless for calculation of attenuation rates at extremely-low-frequencies (less than 1 kc, say), the exponential models find direct application. For example, if $Q=1$ the homogeneous ionosphere is regained and the attenuation rate $(-k \text{ Im } S_0)$ is proportional to $\omega^{\frac{1}{2}}$; on the other hand for an exponential profile with $N(z)$ increasing upwards, the attenuation rate increases with frequency more rapidly as suggested by experimental data. In fact, it is quite easy to see, for both decreasing and increasing exponential profiles, that

$$\frac{\text{Attenuation rate for exponential model}}{\text{Attenuation rate for homogeneous model}}=\sqrt{2}\,|Q|\sin\left(\frac{\pi}{4}-\arg Q\right).$$

TABLE 1. *Selected numerical values of the factor Q*

	$Q=\frac{I_0(\sqrt{i}\,x)}{I_1(\sqrt{i}\,x)}$		$Q=\frac{K_0(\sqrt{i}\,x)}{K_1(\sqrt{i}\,x)}$	
x	$\|Q\|$	arg Q	$\|Q\|$	arg Q
0.2	10.000	−44°43′	0.3854	23°58′
0.5	4.003	−43°13′	0.5880	17°16′
1.0	2.026	−37°55′	0.7344	11°59′
1.5	1.417	−29°52′	0.8047	9°15′
2.0	1.180	−20°59′	0.8459	7°28′
2.5	1.100	−13°44′	0.8728	6°20′
3.0	1.083	−9°09′	0.8919	5°28′
4.0	1.080	−5°42′	0.9169	4°18′
5.0	1.073	−4°37′	0.9325	3°33′
6.0	1.060	−3°49′	0.9433	3°02′
7.0	1.051	−3°19′	0.9511	2°37′
8.0	1.045	−2°47′	0.9570	2°16′
9.0	1.040	−2°26′	0.9618	2°06′
10.0	1.035	−2°11′	0.9654	1°54′
0.1			0.25518	27°22′
0.3			0.4719	21°08′
0.7			0.6622	14°38′

The right-hand side of the above equation is a function of the quantity $x=|\overline{N}|l\cong(\overline{\omega}_r/\omega)^{\frac{1}{2}}l$ where $\overline{\omega}=\overline{\sigma}/\epsilon_0$ is the conductivity at the lower edge of the ionosphere and l is the vertical distance in which $\overline{N}$ changes by a factor 2.718.

The author thanks Mr. Kenneth Spies for his careful reading of the manuscript and Mrs. Ann Fails for her typing and assistance in preparing the bibliography.

14. References

[1] G. N. Watson. The transmission of electric waves round the earth, Proc. Roy. Soc. **95**, 546 (1919).
[2] O. E. H. Rydbeck. On the propagation of radio waves, Trans. Chalmers Univ., **34** (1944).
[3] H. Bremmer. Terrestrial radio waves (Elsevier Publishing Co., Amsterdam, Netherlands, 1949).
[4] W. O. Schumann. Uber die ausbreitung sehr langer elektrischer wellen um die erde und die signale des blitze, II Nuovo cimento **IX** (1952).
[5] W. O. Schumann. Uber die oberfelder bei der Ausbreitung langer, elektrischer wellen im system erde-luft-ionosphare und 2 anwendungen (horizontaler und senkrechter dipol) Z. angew. Phys. **6,** 34, (1954).
[6] W. O. Schumann. Uber die strahlung langer wellen des horizontalen dipols in dem lufthohlraum zwischen erde und ionosphare, Z. angew. Phys. **6,** 225, (1954).
[7] H. Kaden. Die reflexions und schirmwirkung metallischer hullen in einer ebenen elektromagnetischen welle, Sonderdruck aus Archiv der Elektrischen Uberstragung, 403 (1957).
[8] W. L. Taylor and L. J. Lange. Some characteristics of vlf propagation using atmospheric waveforms (Proc. 2d Conf. on Recent advances in atmospheric electricity, Pergamon Press, New York, N.Y., 609, 1959).
[9] F. W. Chapman and R. C. V. Macario. Propagation of audio frequency radio waves to great distances, Nature, **177,** 930 (1956).
[10] S. Namba. General theory of the propagation in the upper atmosphere, Proc. IRE **21,** 238 (1933).
[11] K. G. Budden. The propagation of a radio atmospheric II, Phil. Mag. **43,** 1179 (1952).
[12] K. G. Budden. The propagation of very low frequency radio waves to great distances, Phil. Mag., **44,** 504 (1953).
[13] K. G. Budden. The "wave-guide mode" theory of the propagation of very-low-frequency radio waves, Proc. IRE, **45,** 772 (1957).
[14] IA. L. Al'pert. Low frequency electromagnetic wave propagation over the surface of the earth (Monograph published by Academy of Science, U.S.S.R. 1955).
[15] IA. L. Al'pert. The field of long and very long radio waves over the earth under actual conditions, Radiotekh. i Elektron., **1,** 281 (1956).
[16] L. Lieberman. Anomalous propagation below 500 cps, Proc. Symp. on Propagation of vlf Radio Waves **3,** paper 25, Boulder, Colo., (1957).
[17] J. R. Wait and L. B. Perry. Calculations of ionospheric reflection coefficients at very low radio frequencies J. Geophys. Research **62,,** 43 (1957).
[18] J. R. Wait. On the mode theory of vlf ionospheric propagation, Rev. Geofis. pura e appl. **37,** 103 (1957).

[19] J. R. Wait and H. H. Howe. The waveguide mode theory of vlf ionospheric propagation, Proc. IRE **45,** (1957).
[20] J. R. Wait. The mode theory of vlf ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760 (1957).
[21] J. R. Wait. An extension to the mode theory of vlf ionospheric region, J. Geophys. Research **63,** 125 (1958).
[22] J. R. Wait. Propagation of very-low-frequency pulses to great distances, J. Research, NBS **61,** 187 (1958).
[23] H. H. Howe and J. R. Wait. Mode calculations for vlf ionospheric propagation (Vlf Symp. paper 36, Boulder, Colo., 1957).
[24] B. Friedman. Low frequency propagation in the ionosphere, (Proc. Symp. on the propagation of radio waves, Liege, Belgium, 1958).
[25] A. N. Sommerfeld. Uber die ausbreitung der wellen in der drachtlosen telegraphie, Ann. Physik **81,** 1135 (1926).
[26] A. N. Sommerfeld. Partial differential equations (Academic Press, New York, N.Y., 1949).
[27] F. Magnus and F. Oberhettinger. Special functions of mathematical physics (Chelsea Pub. Co., New York, N.Y., 71, 1949).
[28] W. O. Schumann. Uber elektrische eigenschwingungen des hohlraumes erde-luft-ionosphere are arregt durch blitzentladungen, Z. angew. Phys. **9,** 373 (1957).
[29] J. R. Wait. Low frequency radiation from a horizontal antenna over a spherical earth, Can. J. Phys. **34,** 586 (1956).
[30] K. G. Budden. The reflection of very low frequency radio waves at the surface of a sharply bounded ionosphere with superimposed magnetic field, Phil. Mag. **44,** 504 (1951).
[31] J. R. Wait. Transmission and reflection of electromagnetic waves in the presence of stratified media, J. Research NBS **61,** 205 (1958).
[32] J. R. Wait. Oblique reflection of radio waves from an inhomogeneous region, J. Appl. Phys. **23,** 1403 (1952).

15. Additional References

Al'pert, IA. L., On the speed of propagation of electromagnetic waves at audio frequencies, J. Exp. Theoret. Phys. (U.S.S.R.) **33,** 1305 (1957).
Al'pert, IA. L., Investigation of the propagation of long and very long radio waves by the method of analyzing the shape of atmospherics, Radiotekh. i Elektron., **1,** 293 (1956).
Al'pert, IA. L., and Bordina, S. V., On the velocity of propagation of audio frequency electromagentic waves, Radiotekh. i Elektron. **4,** 195 (1959).
Anderson, C. N., Correlation of long wave transatlantic radio with other factors affected by solar activity, Proc. IRE **16,** 207 (1928).
Anderson, C. N., Notes on radio transmission, Proc. IRE **19,** 1150 (1931).
Austin, L. W., A series of reports on long-distance receiving measurements at Washington, Proc. IRE **12–19** (1931).
Austin, L. W., Preliminary note on proposed changes in the constants of the Austin-Cohen transmission formula, Proc. IRE **14,** 377 (1926).
Bracewell, R. N., Budden, K. G., Ratcliffe, J. A., Straker, T. W., and Weekes, K., The ionospheric propagation of low and very-low frequency radio waves over distances less than 1,000 km, Proc. Inst. Elec. Engrs. **98,** 221 (1951).
Bremmer, H., Propagation of electromagnetic waves, Handbuch der Physik **16,** 423 (1958).
Bremmer, H., Mode expansion in the low frequency range for propagation through a curved stratified atmosphere, J. Research NBS **63D,** 75 (1959).
Brown, J. N., Round-the-world signals at very low frequency, J. Geophys. Research **54,** 367 (1949).
Budden, K. G., The numerical solution of the differential equations governing the reflexion of long radio waves from the ionosphere. II, Phil. Trans. Roy. Soc. London, Sec. A. **248,** 45 (1955).
Budden, K. G., The numerical solution of differential equations governing reflexion of long radio waves from the ionosphere. I, Proc. Roy. Soc. **227,** 516 (1955).
Budden, K. G., A reciprocity theorem on the propagation of radio waves via the ionosphere, Phil. Mag., **44,** 604 (1953, 1954).
Crombie, D. D., Differences between east-west and west-east propagation of vlf signals over long distances, J. Atmospheric and Terrest. Phys. **12,** 110 (1958).
Eckersley, T. L., Studies in radio transmission, J. Inst. Elec. Engrs., **71,** 405 (1932).
Friedman, B. F., Propagation in a nonhomogeneous atmosphere, (Theory of Electromagnetic Waves, Academic Press, New York, N.Y., 1951).
Helliwell, R. A., and Gehrels, E., Observations of magneto-ionic duct propagation using man-made signals of very low frequency, Proc. IRE **46,** 785, (1958).

Heritage, J. L., Weisbrod, S., and Bickel, J. E., A study of signal vs. distance data at vlf, Vlf Symp. Record, **4,** 77 (1957).
Hollingworth, J., The propagation of radio waves, J. Inst. Elec. Engrs., **64,** 579 (1926).
Jean, A. G., Lange, L. J., and Wait, J. R., Ionospheric reflection coefficients at vlf from sferics measurements, Geofis. pura e appl. **38,** 147 (1957).
Karbowiak, A. E., Guided wave propagation in submillimetric region, Proc. IRE **46,** 1706 (1958).
Kendrick, G. W., Radio transmission formulae, Phys. Rev. **31,** 1040 (1928).
Kitchen, F. A., Pressey, B. G., and Tremellen, K. W., A review of present knowledge on the ionospheric propagation of very-low, low- and medium-frequency waves, Proc. Inst. Elec. Engrs. **100,** 100 (1953).
Norton, K. A., Low and medium frequency radio propagation, (Proc. Symp. on propagation of radio waves, Liege, Belgium, 1958).
Norton, K. A., The calculation of the ground wave field intensity, Proc. IRE **29,** 723 (1941).
Pierce, J. A., Sky wave field intensity 1: Low- and very-low radio frequencies, Cruft Lab. Tech. Rept. 158, Harvard Univ. (1952).
Rosner, O., Die ausbreitung elektromagnetischer TE-(H)-wellen, erzeugt durch eine horizontal strom durchflossene leiterscheife, Z. angew. Phys. **9,** 448 (1957).
Round, H. J., Eckersley, T. L., Tremellen, K., and Lunnon F. C., Report on measurements made on signal strength at great distances during 1922 and 1923 by an expedition sent to Australia, J. Inst. Elec. Engrs. **63,** 933 (1925).
Schmelovsky, K. H., Problem der ausbreitung in tropospherischen und ionosphischen wellenleiter, Abhanglungen des Meteorologischen und Hydrologischen Dienstes der Deutschen Demokratischen Republik, **7** (1958).
Schumann, W. O., Uber wellenausbrietung im plasma zwischen zwei unendlich gut leitenden ebenen in richtung eines aufgepragten auberen magnetfeldes, Z. angew. Phys. **8,** 482 (1956).
Schumann, W. O., Uber die zeitliche form und das spehtrum ausqesendeter dipolsignale in einer dielectrischen hohlkugel mit leitenden wanden (Bavarian Academy of Sciences, Munich, Germany, 1956).
Storey, L. R. O., An investigation of whistling atmospherics, Phil. Trans. Roy. Soc., **246,** 113 (1953).
Tremellen, K., A study of ionospheric ray field intensities in the 10–30 and 500–1,100 kc bands, Marconi Rev. **13,** 153 (1950).
Wait, J. R., The attenuation versus frequency characteristics of vlf radio waves, Proc. IRE **45,** 768 (1957).
Wait, J. R., Transmission loss curves for propagation at very low radio frequencies, IRE Trans. on Commun. Systems, **CS–6,** 58 (1958).
Wait, J. R., and Howe, H. H., Amplitude and phase curves for ground-wave propagation in the band 200 cps to 500 kc NBS Circ. **574** (1956).
Wait, J. R., Diffraction corrections to the geometrical optics of low frequency propagation (Proc. Symp. on propagation of radio waves, Liege, Belgium, 1958).
Wait, J. R., Radiation from a vertical antenna over a curved stratified ground, J. Research NBS **56,** 237 (1956).
Wait, J. R., Note on the theory of radio propagation over an ice covered sea, Def. Research Tele. Est., Radio Physics Lab., Project Rept. 18–0–7 (1954). (Gives solution for dielectric slab over a conductive base.)
Wait, J. R., and Murphy, A., Multiple reflections between the earth and the ionosphere in vlf propagation, Geofis. pura e appl. **35,** 61, (1956).
Wait, J. R., Reflection of electromagnetic waves obliquely from an inhomogeneous medium, J. Appl. Phys. **23,** 1403 (1952).
Wait, J. R., and Perry, L. B., Calculations of ionospheric reflection coefficients at very low radio frequencies, J. Geophys. Research **62,** 43 (1957).
Wait, J. R., Mixed path ground wave propagation: 1. Short distances, J. Research NBS **57** (1956).
Wait, J. R., Amplitude and phase of the low-frequency ground wave near a coastline, J. Research NBS **58,** 237 (1957).
Wait, J. R., and Householder, James, Mixed-path ground-wave propagation: 2. Larger distances, J. Research NBS **59,** 19 (1957).
Watt, A. D., and Maxwell, E. L., Characteristics of atmospheric noise from 1 to 100 kc, Proc. IRE **45,** 787 (1957).
Weidner, Julius, Uber eine verfeinerung der theorie der ausbreitung langer elektrischer wellen um die erde z. angew. Phys. **7,** 77 (1955).
Wilkes, M. V., The theory of reflection of very long wireless waves from the ionosphere, Proc. Roy. Soc. **A175,** 143 (1940).
Wilkes, M. V., The oblique reflection of very long wireless waves from the ionosphere, Proc. Roy. Soc. **A189,** 130 (1947).
Yokoyama, E., and Nakai, T., East-west and north-south attenuations of long radio waves in the Pacific, Proc. IRE **17,** 1240 (1929).
Yokoyama, E., and Nakai, T., Effect of sunspots and magnetism on long distance reception on vlf waves Proc. IRE **19,** 882 (1931).
Yokoyama, E., and Tanimua, I., Some long distance transmission phenomena of low frequency waves, Proc. IRE, **21** 263 (1933).

The following two papers have just appeared:

Barron, D. W., The "waveguide mode" theory of radio wave propagation when the ionosphere is not sharply bounded, Phil. Mag. **50,** 1068 (1959). (This is a continuation of early unpublished work of Budden.)

Poverlein, H., Lang-und Langstwellausbreitung, Fortschr. der Hochfrequenztech. **4,** 47 (1959). (Many additional references are found in this excellent review article.)

BOULDER, Colo. (Paper 64D2–49)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 2, March–April 1960

Aurora of October 22/23, 1958, at Rapid City, South Dakota

F. E. Roach* and E. Marovich*

(September 8, 1959)

During the night of October 22/23, 1958, auroral activity at Rapid City, South Dakota, included (a) a visible aurora in the northern part of the sky and; (b) a "monochromatic" (6300 A) arc through the zenith with an azimuth 12° from east-west (geomagnetic). The intensity changes of the arc were independent of the changes in the visible aurora. It moved slowly southward during the night corresponding to a linear speed of about 8 meters per second if its height is 300 kilometers. It is suggested that it is a member of a family of monochromatic arcs which have until recently escaped detection because their red color makes them invisible even though intrinsically intense.

1. Introduction

In an earlier paper [1][1] we described a "monochromatic" (6300 A) arc at Fritz Peak.[2] It was detached from and apparently independent of a visible aurora to the north.

According to Barbier [2] who has made a systematic study of 6300 A auroras at Haute Provence, they occur during about 10 percent of the time either as a general enhancement near the northern (geomagnetic) horizon or as arcs with maxima in the eastern and western parts of the sky. In searching our records at both Fritz Peak and Rapid City, we find, in agreement with Barbier, that they are not infrequent. In this paper, we describe an aurora at Rapid City (longitude 103°18′50″ W; latitude 44°01′ N; geomagnetic latitude 53° N) on October 22/23, 1958, which had a well-defined 6300 A arc near its southern edge.

2. Observational Material

During the night of October 22/23, 1958, the Rapid City photometer was making systematic sky surveys in the three colors—5577, 6300, and 5893 A. The surveys were in the form of zenith observations plus almucantar sweeps at zenith distances of 80°, 75°, 70°, 60°, and 40°. The complete cycle for the three colors required 15 min. The spectral discrimination was by means of a birefringent filter [3]

In figure 1 (left) is shown an example of a set of the original records. One of the striking features is the existence of sharp maxima or spikes on the 6300 A records near the east and west points of each of the sweeps, which are absent from the 5577 A records.[3] When an isophote map is made of all the observations (right side of fig. 1) the spikes are seen to define a 6300 A arc extending across the sky in a general east-west direction. For both colors, there is activity near the northern horizon which developed into a bright aurora during the night.

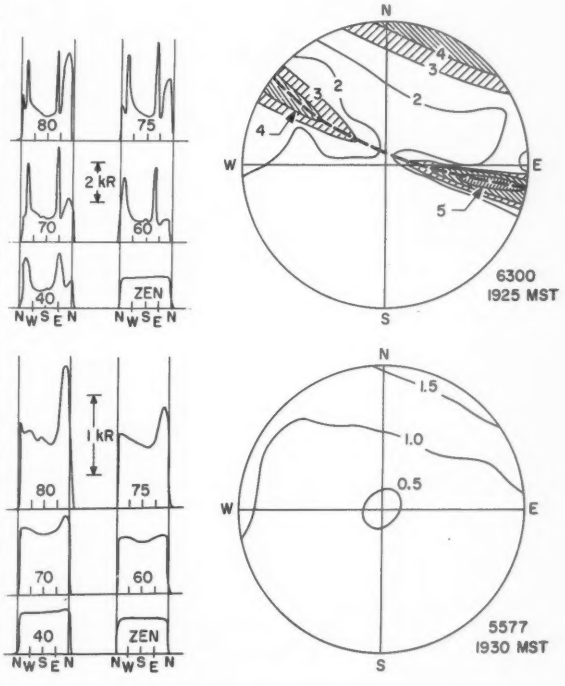


FIGURE 1. *The original records (left) for 6300 A and 5577 A at Rapid City for 1925 and 1930 m.s.t. on the night of October 22, 1958.*

Isophote maps of the sky (right) based on the original records. In drawing the isophote maps a height of 100 km was assumed for 5577 A; of 300 km for 6300 A. The unit of intensity is the kilorayleigh.

*Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[1] Figures in brackets indicate the literature references at the end of this paper.

[2] In this paper the word "monochromatic" is used to indicate that the 6300 A radiation is predominant over the 5577 A.

[3] It is estimated that if 5577 A spikes as weak as 100 R were present, they would be detectable.

3. Evolution of the 6300 A Arc

In figure 2 a pair of isophote maps shows a comparison of 5577 and 6300 A over the sky later in the night when the visible aurora was well developed in the northern part of the sky. The 6300 A arc is seen to have maintained its identity and discreteness independent of the visible aurora.

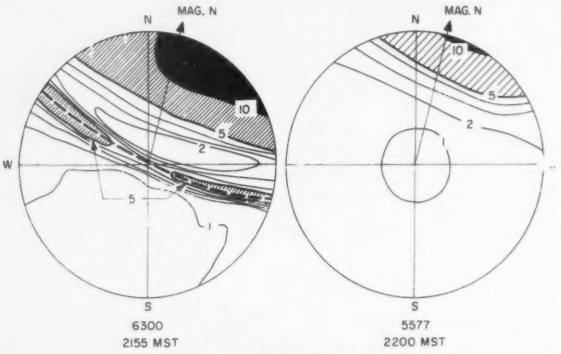


FIGURE 2. *Isophote maps for 6300 A and 5577 A for 2155 and 2200 m.s.t. after the development of an aurora over the northern part of the sky.*

To illustrate the independence of the arc, we show in figure 3 the time variations of 6300 A intensity; (a) in the extreme north-northeast for the visible aurora and; (b) for the 75° E and W spikes of the arc. It is seen that the arc intensity seems unrelated to that of the northern aurora going steadily down, as the northern aurora goes through strong variations, often as much as 10 times the brightness of the arc.

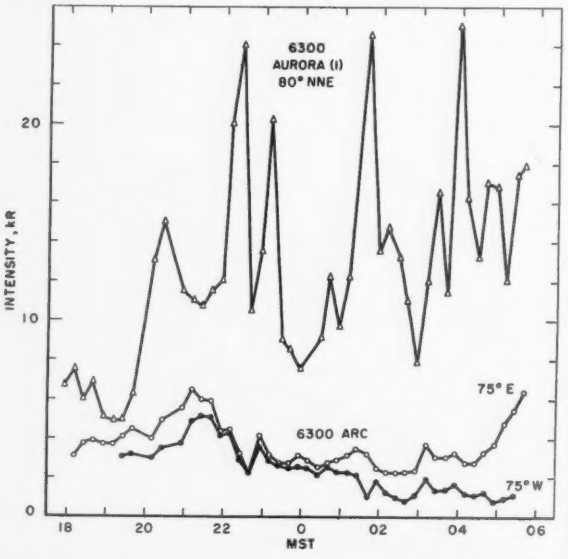


FIGURE 3. *Comparison of 6300 A time variations in two parts of the sky—at 80° zenith distance in the N-NE and the spike in the eastern and western parts of the sky at z=75°.*

In figure 4 we show the position of the 6300 A arc for four different times during the night. It moves steadily southward, as shown also in figure 5. From figure 5 we deduce a mean speed of 8 m/sec for an assumed height of 300 km. This compares with 6 m/sec for the similar arc of September 29/30, 1957, at Fritz Peak [1]; 58 m/sec for the 6300 A front at Sacramento Peak on February 10/11, 1958 (Manring and Pettit [4]); and 100 m/sec for 5577 A airglow cells [5].

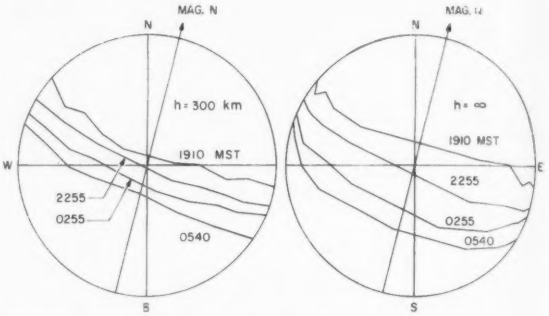


FIGURE 4. *Movement of the 6300 A arc across the sky—as observed in the sky (right, h = ∞) and for an assumed height of 300 km (left).*

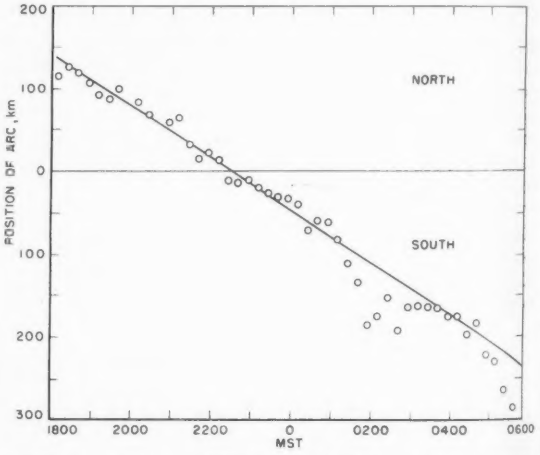


FIGURE 5. *Position of 6300 A arc plotted against time for an assumed height of 300 km.*

4. Tilt of the Arc

In figures 1 and 2 we note that the arc seems to increase in intensity from the zenith to the horizon reminiscent of the so-called van Rhijn effect in the night airglow. This effect reverses during the night, however, as seen in figure 6. The reversal occurs between 2210 and 2225 m.s.t. Referring to figure 7 we note that this is the interval during which the arc crosses the zenith. We thus have the empirical facts that:

(a) When the arc is north of the zenith, it increases in brightness toward the horizon.

(b) When the arc moves south of the zenith, it decreases in brightness toward the horizon.

This combination of facts is consistent with the assumption that the arc is inclined to the vertical along the magnetic lines of force (fig. 8). When the arc is to the north of the zenith, the observer tends to look *across* the arc, which favors a higher intensity toward the horizon. As the arc moves to the south of the zenith, the observer looks more and more *along* the arc favoring a larger effective brightness near the zenith.

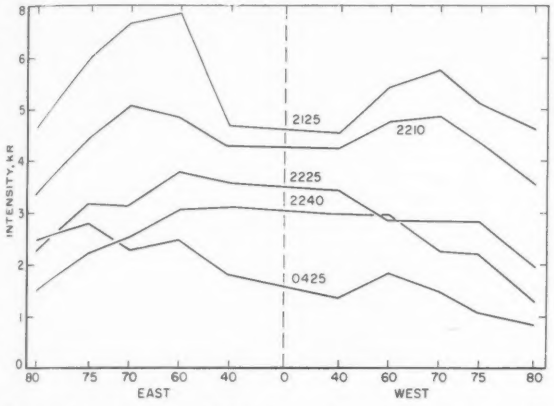


FIGURE 6. *Intensity of 6300 A arc along its length as seen by the observer on the ground at Rapid City for several times during the night of October 22/23, 1958.*

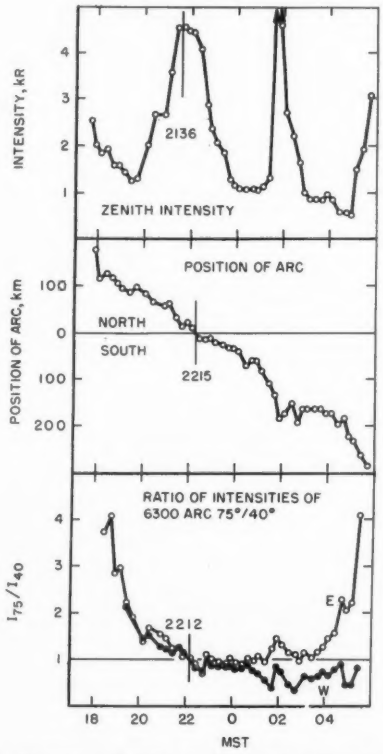


FIGURE 7. *Changes with time of: (a) zenith intensity of 6300 A, (b) position of the 6300 A arc, and, (c) the ratio* I_{75}/I_{40}.

5. Apparent Curvature of the Arc

Figures 1, 2, and 4 all show that our method of plotting results in a curved arc which can be explained by assuming that the arc follows a parallel of latitude (neither geographic nor geomagnetic, as we shall see in the next section). The amount of curvature should, in principle, give a value of the effective height of the arc, but very precise observations would be required to yield a reliable height by this method. A comparison of the observed curvature with that expected for assumed heights of 100, 200, and 300 km indicates that the height is at least 300 km. Throughout this paper, we have assumed an effective height of 300 km.

6. Orientation of the Arc

An inspection of figures 1, 2, and 4 shows clearly that the arc is oriented along the west-northwest to east-southeast direction. This is only *approximately* at right angles to magnetic north. Magnetic north at Rapid City is about 14° east of geographic north; geomagnetic north is 11°.3 east of geographic north. We have estimated on working plots the direction of the perpendicular to each arc and find the following:

Mean of 41 measures	22°.9 east of geographic north
Extremes	24°.7 and 20°.2.

Thus, the observed arc is inclined to magnetic parallels by an angle of about 9° and to geomagnetic parallels by an angle of about 12° in the sense that the western part is north and the eastern part south of the parallels.

It is of interest to note that this behavior is similar to that observed for polar auroral arcs. For example, Vegard and Krogness [6] and Vegard [7] found an inclination of 10° for arctic data from the first polar year. Jensen and Currie [8] report a mean inclination of 6° for observations made in Canada. In both cases the sense of the inclination is the same as we have noted in the October 22/23 arc at Rapid City.

7. Southward Movement

The hypothesis might be advanced that the southward movement of the arc during the night is due to the rotation of the observer on the earth's surface under a fixed arc oriented in a west-northwest, east-southeast direction. Calculation shows that a small angle of inclination (between 1° and 2°) is sufficient to account for the observed movement of about 8 m/sec. Since the observed angle is much larger (23°), the hypothesis is not supported.

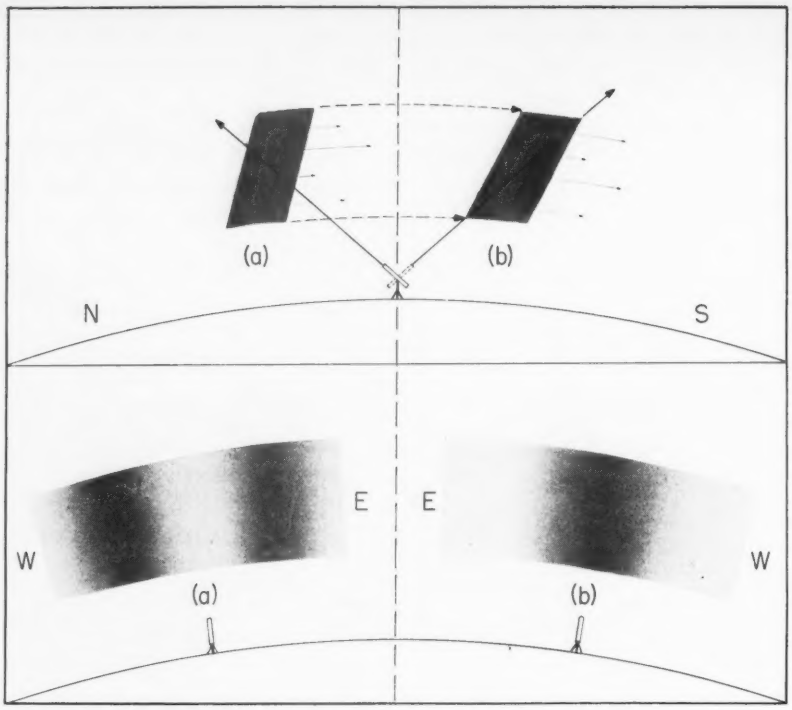


FIGURE 8. *Section (above) and plan (below) views of the arc; (a) to the north of the observer and; (b) to his south.*

The shading in the plan view corresponds to the observed zenith intensity dip when the arc is to the north (a) and the zenith intensity maximum when the arc moves to the south.

8. Discussion

The evidence of this paper, of our earlier study [1], and of the work of Barbier at Haute Provence [2] suggests that there is a family of low latitude auroras with the following characteristics:

(a) The features are "monochromatic" (6300 A present; 5577 A absent). This implies that the excitation mechanism is sufficient to excite the oxygen atom to its ^1D state (2.0 ev) but not to the ^1S state (4.2 ev).

(b) They occur at geomagnetic latitudes as low as 48°. The fact that they have not been observed *overhead* at Haute Provence (Φ=46°) suggests that 48° is actually the lower limit.

(c) They occur as arcs slightly inclined to the perpendicular to the earth's magnetic field.

(d) They may occur concurrently with visible auroras but they seem to be independent of the visible auroral activity both with regard to geographical position and absolute intensity.

(e) 6300 A arcs have been observed in the southern hemisphere at Camden, Australia (Φ=42° S) by Duncan [9]. On September 5, 1958, the apex of such an arc was 74° to the geomagnetic south. If the height is 300 km, the arc would be 820 km or 7°.5 to the south, placing it at a geomagnetic latitude, Φ, of 49°.5.

It is possible that 6300 A arcs are a general feature on the equatorial side of all auroras. If so, they would have escaped general detection since the visual threshold in the red is probably between 6 and 10 kilorayleighs. The two arcs already studied by us had maximum intensities of 6 to 7 kR and were detected only by means of a photometer and not visually. The *color* threshold for 6300 A is probably about 30 kR. We can therefore summarize the nature of the possible observations of such arcs.

Class	Absolute intensity (kR)	Nature of observation
1	1 to 6	Not visible; detectable only with photometers
2	6 to 30	Visible but no color
3	>30	Visible as colored arcs

Most of the historical data on auroral occurrence are based on visual observations, so class 1 arcs (intensity $<$ 6 kR) which might be expected to be the most frequent are just now being studied with monochromatic photometers.

It is possible that class 2 arcs (intensity 6 to 30 kR) may be included in a compilation by Störmer [10]. Many of the features of certain arcs described

by Störmer suggest similarities with the arc described in this paper. It is not possible to deduce whether Störmer's arcs were monochromatic (6300 A) or even approximately so. His observations are concentrated over southern Norway at $\Phi=56°$.

The brighter class 3 arcs which appear red (6300 A) to the eye (intensity $>$ 30 kR) seem to be very rare. Störmer [11] records only two such arcs over southern Norway between 1911 and 1940. In a historical review, he lists nine between 1826 and 1872 and concludes "that red arcs very rarely occur, and only during years of great solar activity."

In our earlier paper [1] we suggested the possibility of some relationship between the 6300 A arcs and the newly discovered radiation belts. The arc discussed in the present paper corresponds to a dipole line of force cutting the magnetic equator at a distance of 2.8 earth radii from the center of the earth, just inside the inner edge of the outer radiation belt [12].

The authors are especially indebted to J. E. Cruz who was in charge of the observations at Rapid City during the International Geophysical Year and on the particular night under analysis in this paper. We have profited also by stimulating discussions with D. R. Bates, J. W. Chamberlain, and D. M. Hunten during a stay at the Yerkes Observatory during July–August 1959. The study was supported in part by the National Science Foundation during the I.G.Y. The visit to the Yerkes Observatory where the final discussion of the material was completed was sponsored by the Air Force Cambridge Research Center under Contract AF 19(604)–3044.

9. References

[1] F. E. Roach and E. Marovich, A monochromatic low latitude aurora, J. Research NBS **63D,** 297 (1959)
[2] D. Barbier, L'activite aurorale aux basses latitudes, Ann. Geophys. **14,** 334 (1958)
[3] R. B. Dunn and E. R. Manring, Recording night sky photometer of high spectral purity, J. Opt. Soc. Am. **46(8),** 572 (1956)
[4] E. R. Manring and H. B. Pettit, Photometric observations of the 5,577 A and 6,300 A emissions made during the aurora of February 10–11, 1958, J. Geophys. Research **64,** 149 (1959)
[5] F. E. Roach, E. Tandberg-Hanssen, and L. R. Megill, Movements of airglow cells, J. Atmospheric and Terrest. Phys. **13,** 122 (1958)
[6] L. Vegard and O. Krogness, The position in space of the aurora polaris from observations made at the Haldde Observatory, 1913–1914, Geofys. Publikasjoner **1,** 1 (1920)
[7] L. Vegard, Recent results of northlight investigations and the nature of the cosmic electric rays, Phil. Mag. **42,** 47 (1921)
[8] R. E. Jensen and B. W. Currie, Orientations of auroral displays in west-central Canada, J. Geophys. Research **46,** 269 (1953)
[9] R. A. Duncan, Photometric observations of subvisual red auroral arcs at middle latitudes, Australian J. Phys. **12,** No. 2 (1959)
[10] Carl Störmer, Remarkable aurora—forms from southern Norway, I. Feeble homogeneous arcs of great altitude, Geofys. Publikasjoner **11,** No. 5 (1934–37)
[11] Carl Störmer, Remarkable aurora—forms from southern Norway, VII. Red patches and red arcs, VIII. Some statistics on red aurorae in Norway, Geofys. Publikasjoner **13,** 57 (1940–1944)
[12] J. Van Allen, C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958 epsilon, J. Geophys. Research **64,** 271 (1959)

Boulder, Colo. (Paper 64D2–50)

Selected Abstracts

On a modification of Watson's lemma, F. Oberhettinger, *J. Research NBS* **63B,** *No. 1, 15 (1959).*

The method of steepest descents is extended to the case when a saddle point and a pole of arbitrary order are involved. An application to a problem in diffraction theory is demonstrated.

Frequency dependence of VHF ionospheric scattering, J. C. Blair, *NBS Tech. Note 9 (PB151368) 75 cents.*

First results are given from an improved experimental study of frequency dependence of vhf ionospheric scatter transmission. Continuous recording of transmission loss was carried out simultaneously at five frequencies from 30 to 108 Mc, using narrow beam antenna systems, scaled with frequency for constant gain and pattern. Hourly median values of transmission loss were observed to be proportional to frequency raised to power n_s; n_s varied diurnally and seasonally, with typical values obtained from 7.3 to 9.9 in March 1958 and 6.8 to 8.2 in June 1958. (The exponent associated with frequency dependence of scattering cross section σ is n_s-2.) Dependence of signal fading rate and amplitude distribution on carrier frequency was also observed. Differential effects of absorption and scattering are noted during solar flares, depending on frequency.

Precise time synchronization of widely separated clocks, A. H. Morgan, *NBS Tech. Note 22 (PB151381) $1.50.*

The note describes several methods of setting a group of widely separated clocks to the same time and keeping them in close agreement Particular attention is focused on the use of high-frequency radio waves as a link. Using a simplified theory of propagation the influence of the variation of the heights of the ionospheric *F2* layer is discussed. This layer is assumed to be perfectly conducting, and all refraction effects are neglected. The time delay of the ground wave is assumed not to be influenced by earth curvature and finite conductivity.
The note contains a valuable set of curves which plot the time delay Δt as a function of range for various sky-wave hops. Δt is described in the note as "the interval between successive hops." Actually as plotted, Δt is the difference in length (expressed in microseconds) between the n hop sky wave and the great circle distance measured along the ground. Extensive data used for plotting these curves were computed by Dr. H. Herbert Howe.

Design of single frequency filters, F. F. Fulton, Jr., *NBS Tech. Note 23 (PB151382) 50 cents.*

Efficient procedures are shown for designing filters formed by a number of identical resonant circuits loosely coupled together, and which are required to accept one narrow band of frequencies and reject another narrow band somewhat removed in frequency, without any special requirements on the shape of the attenuation curve in between these regions. The design is based on using a large number of sections.

Radio wave propagation in an inhomogeneous atmosphere, J. R. Wait, *NBS Tech. Note 24 (PB1. 1383) $1.00.*

A self-contained treatment of the theory of radio waves in an inhomogeneous atmosphere is given. The refractive index is assumed to vary with height above the earth in a monotonic fashion. Variation according to an exponential law is used for illustration of general principles. For this case, rigorous series formulas are developed for the distance to the horizon for an elevated point in the atmosphere.

Communication theory aspects of television bandwidth conservation, W. C. Coombs, *NBS Tech. Note 25 (PB151384) 50 cents.*

Advances in information theory are noted which indicate possibilities for reduced television transmission bandwidth and/or improved reliability. A review of television transmission standards is suggested to include consideration of digital transmission of *picture-changes information* ("difference-signal" modulation technique). The possibilities are suggested in relation to current television frequency allocation problems and consideration of assignment of new frequency bands. In the VHF range reduced bandwidth might be achieved, or alternatively, utility of UHF channels might be enhanced by improving the transmission reliability. Straightforward parametric interchanges of bandwidth and signal-to-noise ratio have not been found generally profitable toward either of these goals if full conventional video amplitude information capacity has preserved throughout the system. A real advantage is obtained by redistribution, over the available bandwidth, of information capacity more comparable to the capacity for assimilation by the human observer channel.

Applications of the molecular refractivity in radio meteorology, B. R. Bean and R. M. Gallet, *J. Geophys. Research* **64,** *1439 (1959).*

Consideration of the molecular refractivity has led to a value of the radio refractive index referred to zero altitude or sea level that effectively removes the systematic altitude dependence of air density. The advantages of this new parameter are illustrated by two applications. The first application, the preparation of climatic charts of the radio refractive index, illustrates that this new parameter brings climatic differences into sharp relief. The second application deals with the vertical variation of airmass properties across a classic warm front and clearly delineates airmass differences, thus indicating the utility of this new parameter for synoptic studies.

Flare-associated bursts at 18 Mc/s, C. Warwick and J. W. Warwick, *Proc. Paris Symp. on Radio Astronomy (IAU Symp. No. 9 and URSI Symp. No. 1), July 30 to Aug. 6, 1958, Edited by R. N. Bracewell, Paper No. 37, p. 203 (Stanford Univ. Press, 1959).*

Flare-associated bursts at radio emission at 18 Mc are related to major bursts at metric wavelengths, and to both type III (fast-drift) and type II (slow-drift) bursts observed with sweep-frequency equipment. Intense bursts at 18 Mc, occurring before maximum of the cosmic noise absorption that accompanies the flare, are followed after 2 days by a definite rise in level of geomagnetic activity.

Geomagnetic activity following large solar flares, C. S. Warwick and R. T. Hansen, *J. Atmospheric and Terrest. Phys.* **14,** *287 (1959).*

Large solar flares, selected according to their optical characteristics alone, are followed by significant increase in geomagnetic activity, excepting the period 1950 to 1954, near minimum of solar activity. This result contradicts and possibly explains the conclusion of Watson, that geomagnetic disturbance is unrelated to solar flares.

System loss in radio propagation, K. A. Norton, *Letter Proc. IRE* **47**, *1661 (1959).*

This letter contains in full the recommendation of the Consultative Committee on International Radio relative to the concepts of transmission loss in radio systems studies. This recommendation contains definitions of the terms "system loss," "transmission loss," "basic transmission loss," "propagation loss," "path antenna directivity gain," and "path antenna power gain."

Factors affecting modulation techniques for VHF scatter systems, J. W. Koch, *IRE Trans. Commun. Systems* **CS–7**, *No. 2, 77 (1959).*

An experimental program at approximately 50 Mc has been carried out over a 1295-km east-west ionospheric-scatter path to determine the communication capacity of the propagation medium using ordinary modulation techniques.

Observations of the ionosphere over the south geographic pole, R. W. Knecht, *J. Geophys. Research* **64**, *No. 9, 1243 (1959).*

Monthly median values of penetration frequencies of the ionosphere over the south geographic pole have been examined. Twenty months of data (June 1957 through January 1959) are included in the study of f_0F_2, while f_0E and f_0F_1 median values are shown for 6 summer months (November and December 1957; January, November, and December 1958; January 1959). It is found that F-region ionization persists throughout the 6-month winter night. Marked diurnal variations are observed in the monthly medians f_0F_2 even though the usual daily variation in solar elevation is absent at this unique location. A small but significant diurnal variation is also found in f_0F_1. In contrast, f_0E exhibits no regular daily fluctuation, but seems to depend to a greater extent on the level of solar activity.

Other NBS Publications

Journal of Research, Section 64A, No. 1, January–February 1960. 70 cents.

The "1958 He^4 scale of temperatures." F. G. Brickwedde, H. van Dijk, M. Durieux, J. R. Clement, and J. K. Logan.
Energy levels and spectrum of neutral helium (^4He I). W. C. Martin.
Vibration-rotation structure in absorption bands for the calibration of spectrometers from 2 to 16 microns. E. K. Plyler, A. Danti, L. R. Blaine, and E. D. Tidwell.
Combustion calorimetry with fluorine: constant pressure flame calorimetry. G. T. Armstrong and R. S. Jessup.
Pits in metals caused by collision with liquid drops and rigid steel spheres. O. G. Engel.
Theory of formation of polymer crystals with folded chains in dilute solution. J. I. Lauritzen, Jr., and J. D. Hoffman.
Studies of beryllium chromite and other beryllia compounds with R_2O_3 oxides. C. E. Weir and A. Van Valkenburg.
The uranium-platinum system. J. J. Park and D. P. Fickle.
Nitriding phenomena in titanium and the 6Al–4V titanium alloy. J. R. Cuthill, W. D. Hayes, and R. E. Seebold.
Stability of silver and Pyrex in perchloric acid—silver perchlorate solutions and in conductivity water. D. N. Craig, C. A. Law, and W. J. Hamer.
Note on the preparation of sodium amalgam in the form of pellets. H. S. Isbell, H. L. Frush, and N. B. Holt.

Journal of Research, Section 64B, No. 1, January–March 1960. 75 cents.

Theory of diffraction in microwave interferometry. D. M. Kerns and E. S. Dayhoff.
Some solutions for electromagnetic problems involving spheroidal, spherical, and cylindrical bodies. James R. Wait.
Kantorovich's inequality. Morris Newman.
A symmetric continuous poker model. A. J. Goldman and J. J. Stone.
Moebius function on the lattice of dense subgraphs. R. E. Nettleton and M. S. Green.
The minimum of a certain linear form. Karl Goldberg.
Space of k-commutative matrices. Marvin Marcus and N. A. Khan.
Selected bibliography of statistical literature, 1930 to 1957: I. Correlation and regression theory. Lola S. Deming.
Selected bibliography of statistical literature, 1930 to 1957: II. Time series. Lola S. Deming.

Journal of Research, Section 64C, No.1, January–March 1960. 75 cents.

Power loss and operating temperature of tires. R. D. Stiehler, M. N. Steel, G. G. Richey, J. Mandel, and R. H. Hobbs.
Effects of antioxidants on asphalt durability. B. D. Beitchman.
Temperature stratification in a nonventing liquid helium Dewar. L. E. Scott, R. F. Robbins, D. B. Mann, and B. W. Birmingham.
Expansion engines for hydrogen liquefiers. E. H. Brown.
A statistical chain-ratio method for estimating relative volumes of mail to given destinations. N. C. Severo and A. E. Newman.
Standard free-air chamber for the measurement of low energy X-rays (20 to 100 kilovolts-constant potential). V. H. Ritz.
Transmittance of materials in the far infrared. E. K. Plyler and L. R. Blaine.
Equipment and method for photoelectric determination of image contrast suitable for using square wave targets. F. W. Rosberry.
Formation of silver sulfide in the photographic image during fixation. C. I. Pope.
Capacitor calibration by step-up methods. T. L. Zapf.

Recorder survey: recording surfaces and marking methods, G. Keinath, NBS Circ. 601 (1959) 30 cents.
Energy dissipation by fast electrons, L. V. Spencer, NBS Monograph 1 (1959) 45 cents.
Hydraulic Research in the United States—1959, H. K. Middleton, NBS Misc. Pub. 227 (1959) $1.25.
Analysis of ionospheric vertical soundings for electron density profile data. 1. Facilities for convenient manual reduction of ionograms, J. W. Wright and R. B. Norton, NBS Tech. Note 14 (PB151373) (1959) 50 cents.
Some applications of statistical sampling methods to outgoing letter mail characteristics, N. C. Severo, A. E. Newman, S. M. Young, and M. Zelen, NBS Tech. Note 16 (PB151375) (1959) $2.75.
Radio noise data for the International Geophysical Year July 1, 1957—December 31, 1958, W. Q. Crichlow, C. A. Samson, R. T. Disney, and M. A. Jenkins, NBS Tech. Note 18 (PB151377) (1959) $2.50.
Analysis of ionospheric vertical soundings for electron density profile data. II. Extrapolation of observed electron density profiles above $h_{max}F2$, J. W. Wright, NBS Tech. Note 19 (PB151378) (1959) 50 cents.
Variations of gamma cassiopeiae, S. R. Pottasch, NBS Tech. Note 21 (PB151380) (1959) 75 cents.

Aerodynamic phenomena in stellar atmospheres—A bibliography, NBS Tech. Note 30 (PB151389) (1959) $1.25.

Sugar retention by char, F. G. Carpenter, Proc. 5th Tech. Session on Bone Char 1957, p. 279 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Removal of organic anions by bone char, V. R. Dietz and H. M. Rootare, Proc. 5th Tech. Session on Bone Char 1957, p. 297 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Recent advances in cryogenic engineering, R. B. Jacobs, Am. Rocket Soc. ARS J. **29,** 245 (1959).

Transistorized velocimeter for measuring speed in the sea, C. E. Tschiegg, J. Acoust. Soc. Am. **31,** No. 7, 1038 (1959).

Magnetic study of the frozen products from the nitrogen microwave discharge, B. J. Fontana, J. Chem. Phys. **31,** No. 1, 148 (1959).

The thermal E.M.F. of several thermometric alloys, R. L. Powell and M. D. Bunch. Suppl. Bull. Inst. Intern. Froid (Delft, Holland) Comm. 1, 129 (1958).

Electrical discharge induced luminescence of solids at low temperatures, L. J. Schoen and R. E. Rebbert, J. Mol. Spect. **3,** No. 4, 417 (1959).

The nature, cause and effect of the porosity in electrodeposits, V. An evaluation of the sensitivity of the ferroxyl test, F. Ogburn, D. W. Ernst, and W. H. Roberts, Plating **46,** 1052 (1959).

Introductory remarks, A. V. Astin, Proc. 5th Tech. Session on Bone Char 1957, p. 1 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Analysis of liquid sugars, E. J. McDonald, Proc. 5th Tech. Session on Bone Char 1957, p. 77 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Development of a new test for the abrasion hardness of bone char, F. G. Carpenter, Proc. 5th Tech. Session on Bone Char 1957, p. 99 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Effects of controlled decarbonization on the performance of service synthad, F. W. Schwer, W. V. Loebenstein, and E. P. Barret, Proc. 5th Tech. Session on Bone Char 1957, p. 145 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Some mechanisms of color and ash removal by bone char, A. Gee, Proc. 5th Tech. Session on Bone Char 1957, p. 163 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Survey of variations in use for conducting laboratory—scale column filteration tests, V. R. Deitz, Proc. 5th Tech. Session on Bone Char 1957, p. 237 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Comparison of column decolorization experiments with theory, W. V. Loebenstein, Proc. 5th Tech. Session on Bone Char 1957, p. 253 (Bone Char Research Project, Inc., Charlestown, Mass., 1959).

Spectrum of thin target bremsstrahlung bounded by a forward circular cone, J. H. Hubbell, J. Appl. Phys. **30,** 981 (1959).

Negative atomic ions, H. R. Johnson and F. Rohrlich, J. Chem. Phys. **30,** 1068 (1959).

Inclusion theorems for congruence subgroups, M. Newman and I. Reiner, Trans. Am. Math. Soc. **91,** No. 3, 369 (1959).

The very low-frequency emissions generated in the earth's exosphere, R. M. Gallet, Proc. IRE, **47,** No. 2, 211 (1959).

Reaction of hydrogen atoms with solid oxygen at 20° K, R. Klein and M. D. Scheer, J. Chem. Phys. **31,** 278 (1959).

Application of the Williams-Landel-Ferry equation to silicate glasses, A. B. Bestul, Glastech. Ber. (Frankfurt, Germany) **32 K,** No. VI, 59 (1959).

Calculated patterns of slotted elliptic-cylinder antennae, J. R. Wait and W. E. Mientka, Appl. Sci. Research **B 7,** 449 (1959).

Glow discharge spectra of copper and indium above aqueous solutions, D. E. Couch and A. Brenner, Tech. Notes, J. Electrochem. Soc. **106,** 628 (1959).

Free radical chemistry, J. W. Moyer and A. M. Bass, Chem. and Eng. News, p. 51 (Aug. 24, 1959).

Gaseous heat conduction at low pressures and temperatures, R. J. Corruccini, Vacuum **7–8,** 19 (Pergamon Press Ltd., London, England, 1959).

Discussion of the papers of Messrs. Satterthwaite and Budne, W. J. Youden, Technometrics **1,** No. 2, 157 (1959).

New approach in the theory of satellite orbits, J. P. Vinti, Phys. Rev. Letters **3,** No. 1 (1959).

Production of embossing plates from texture patterns by electroforming methods, J. P. Young and V. A. Lamb, Plating **46,** 1033 (1959).

Method of evaluating the clinical effect of warping a denture: Report of a case, J. B. Woelfel and G. C. Paffenbarger, J. Am. Dental Assoc. **59,** 250 (1959).

The sound transmission loss of some building construction, R. V. Waterhouse, R. D. Berendt, and R. K. Cook, Noise Control **5,** No. 4, 40 (1959).

Cryogenic insulation, R. H. Kropschot, Am. Soc. Heating, Refrig. and Air-Conditioning Engrs. **1,** No. 9, 48 (1959).

End plate modification of X-band TE011 cavity resonators, M. S. Thompson, F. E. Freethey, and D. M. Waters, IRE Trans. on Microw. Theory Tech. **MTT–7,** 388 (1959).

Evolution of amplified waves leading to transition in a boundary layer with zero pressure gradient, P. S. Klebanoff and K. D. Tidstrom, NASA Tech. Note D–195, 1 (1959).

Microwave reflectometer, G. F. Engen and R. W. Beatty, IRE Trans. on Microw. Theory Tech. **MTT–7,** 351 (1959).

Multiple ionization of sodium vapor by electron impact, V. H. Dibeler and R. M. Reese, J. Chem. Phys. **31,** No. 1, 282 (1959).

Penetration and diffusion of X-rays, U. Fano, L. V. Spencer, and M. J. Berger, Encyclopedia of Phys. **38,** No. 2, 660 (1959).

Problems of the experimenter, W. J. Youden, Natl. Conv. Trans. Am. Soc. Qual. Control, p. 41 (1959).

Vibrational intensity distributions in the nitrogen afterglow, U. H. Kurzweg and H. P. Broida, J. Mol. Spect. **3,** No. 4, 388 (1959).

Some characteristics of VLF propagation using atmospheric waveforms, W. L. Taylor and L. J. Lange, Recent Advances in Atmospheric Elec. (Proc. 2d Conf. on Atmospheric Elec.), Portsmouth, N.H., May 20–23, 1958, p. 609 (Pergamon Press, Inc., New York, N.Y., 1958).

A new aid for the rapid determination of absorption corrections by Albrecht's method, D. K. Smith, Acta Cryst. (Copenhagen, Denmark) **12,** Pt. 6, 479 (1959).

Measurements made by matching with known standards, W. J. Youden, W. S. Connor, and N. C. Severo, Technometrics **1,** No. 3, 101 (1959).

Influence of crystallographic orientation on the pitting of iron in distilled water, J. Kruger, J. Electrochem. Soc. **106,** No. 8, 736 (1959).

Evacuated powder insulation for low temperatures, M. M. Fulk, Progress in Cryogenics—1, 65 (Heywood and Co., Ltd., London, England, 1959).

A direct-reading viscometer, M. R. Shafer, Instr. and Control Systems **32,** No. 7, 1044 (1959).

Note on quiet-day vertical cross sections of the ionosphere along 75° W geographic meridian, J. W. Wright, J. Geophys. Research **64,** 1631 (1959).

Effect of atomic tests on radio noise, C. A. Samson, Nature **184,** 538 (1959).

An intermittent-action camera with absolute time calibration, G. Hefley, R. H. Doherty, and E. L. Berger, IRE Wescon Conv. Record **3,** Pt. 6, 129 (1959).

Absolute photometry of the aurora—I. The ionized molecular nitrogen emission and the oxygen green line in the dark atmosphere, M. H. Rees, J. Atmospheric and Terrest. Phys. **14,** 325 (1959).

Absolute photometry of the aurora—II. Molecular nitrogen emission in the sunlit atmosphere, M. H. Rees, J. Atmospheric and Terrest. Phys. **14,** 338 (1959).

An additional lunar influence on equatorial E_3 at Huancayo, R. W. Knecht, Research Note, J. Atmospheric and Terrest. Phys. **14,** 348 (1959).

Interplanetary gas. I. Hydrogen radiation in the night sky, J. C. Brandt and J. W. Chamberlain, Astrophysical J. **130,** 670 (1959).

The nova outburst—I. S. Pottasch, Ann. astrophys. **22,** 297 (1959).

The nova outburst—II. The radiative cooling of an expanding shell, S. Pottasch, Ann. astrophys. **22,** 310 (1959).

The nova outburst—III. The ionization of hydrogen gas by an exciting star, J. Jefferies and S. Pottasch, Ann. astrophys. **22,** 318 (1959).

Atomic weights, E. Wichers, ch. 6. Treatise on Analytical Chemistry, pt. 1. Theory and Practice, vol. 1, sec. B. Application of Chemical Principles, p. 161 (The Interscience Encyclopedia, Inc., New York, N.Y., 1959).

Concept and determination of *p*H, R. G. Bates, ch. 10. Treatise on Analytical Chemistry, pt. 1. Theory and Practice, vol. 1, sec. B. Application of Chemical Principles, p. 361 (The Interscience Encyclopedia, Inc., New York, N. Y., 1959).

Electrode potentials, R. G. Bates, ch. 9. Treatise on Analytical Chemistry, pt. 1. Theory and Practice, vol. 1, sec. B. Application of Chemical Principles, p. 319 (The Interscience Encyclopedia, Inc., New York, N.Y., 1959).

Principles and methods of sampling, W. W. Walton and J. I. Hoffman, ch. 4. Treatise on Analytical Chemistry, pt. 1. Theory and Practice, vol. 1, sec. A. Analytical Chemistry: Its Objectives, Functions and Limitations, p. 67 (The Interscience Encyclopedia, Inc., New York, N.Y., 1959).

System loss in radio-wave propagation, K. A. Norton, Letter to Editor, Proc. IRE **47,** No. 9, 1661 (1959).

The oxide films formed on copper single crystal surfaces in pure water. 1. Nature of the films formed at room temperature, J. Kruger, J. Electrochem. Soc. **106,** No. 10, 847 (1959).

Strain gauge calibration device for extreme temperatures, R. M. McClintock, Rev. Sci. Instr. **30,** No. 8, 715 (1959).

Variation of the thermodynamic ideal temperature in the polystyrene-cyclohexane system, D. McIntyre, J. H. O'Mara, and B. C. Konouck, J. Am. Chem. Soc. **81,** 3498 (1959).

Rapid insertion device for coaxial attenuators, A. Y. Rumfelt and R. J. Como, Rev. Sci. Instr. **30,** No. 8, 687 (1959).

Energy dissipation in standing waves in rectangular basins, G. H. Keulegan, J. Fluid Mech. **6,** pt. 1. 35 (1959).

Trapped energetic radicals, J. L. Franklin and H. P. Broida, Ann. Rev. Phys. Chem. **10,** 145 (1959).

Electrode potentials, p. 319, and Concept and determination of *p*H, p. 361, R. G. Bates, pt. 1, vol. 1, Treatise on Analytical Chemistry, Edited by I. M. Kolthoff and P. J. Elving (Interscience Pub. Co., New York, N.Y., 1959).

Total photoelectric cross sections of copper, molybdenum, silver, tantalum, and gold at 662 kev, W. F. Titus, Phys. Rev. **115,** No. 2, 351 (1959).

Diode in feedback loop makes stable transistor bias circuit, G. F. Montgomery, Electronic Design **7,** 54 (1959).

Strength of synthetic single crystal sapphire and ruby as a function of temperature and orientation, J. B. Wachtman and L. H. Maxwell, J. Am. Ceram. Soc. **42,** No. 9, 432 (1959).

PILOT, the NBS multicomputer system, A. L. Leiner, W. A. Notz, J. L. Smith, and A. Weinberger, Proc. Eastern Joint Computer Conf. (Philadelphia, Pa.) p. 164 (1958).

Investigation of creep behavior of structural joints under cyclic loads and temperatures, L. Mordfin, N. Halsey, and G. E. Greene, NASA Tech. Note D–181, p. 37 (1959).

Dimensional stability of impregnated sole leather, B. H. Fouquet, J. Am. Leather Chemists' Assoc. **LIV,** No. 10, 544 (1959).

Theoretical problems of reliability measurement and prediction, J. R. Rosenblatt, Proc. of the Third Exploratory Conf. on Missile Model Design for Reliability Prediction at White Sands Missile Range (New Mexico) p. 29 (1959).

Accuracy and precision: evaluation and interpretation of analytical data, W. J. Youden, ch. 3. Treatise on Analytical Chemistry, pt. I. Theory and Practice, vol. 1, sec. A. Analytical Chemistry: Its Objectives, Functions and Limitations, p. 47 (The Interscience Encyclopedia, Inc., New York, N.Y., 1959).

Dimensional changes in systems of fibrous macromolecules: polyethylene, L. Mandelkern, D. E. Roberts, A. F. Diorio, and A. S. Posner, J. Am. Chem. Soc. **81,** 4148 (1959).

Microwave reflectometer techniques, G. F. Engen and R. W. Beatty, IRE Trans. on Microw. Theory and Tech. **MTT–7,** No. 3, 346 (1959).

Attenuation of scattered cesium-137 gamma rays, F. S. Frantz, Jr., and H. O. Wyckoff, Radiology **73,** No. 2, 263 (1959).

Dependence of mechanical strength of brittle polycrystaline specimens on porosity and grain size, F. P. Knudsen, J. Am. Ceram. Soc. **42,** No. 8, 376 (1959).

Theory of dielectric relaxation in molecular crystals, J. D. Hoffman, Arch. sci. (Geneva), Soc. Phys. Hist. Nat. Geneva **12,** 36 (1959).

Denture reliners—direct, hard, self-curing resin, G. M. Brauer, E. E. White, Jr., C. L. Burns, and J. B. Woelfel, J. Am. Dental Assoc. **59,** 270 (1959).

Photoproton cross sections of carbon, S. Penner and J. E. Leiss, Phys. Rev. **114,** No. 4, 1101 (1959).

Cryogenic engineering of hydrogen bubble chambers, B. W. Birmingham, D. B. Chelton, D. B. Mann, and H. P. Hernandez, ASTM Bull. No. 240, 34 (TP164) (1959).

A porcelain reference electrode conductive to sodium ions for use in molten salt systems, R. J. Labrie and V. A. Lamb, J. Electrochem. Soc. **106,** No. 10, 895 (1959).

Memorandum on a procedure for obtaining spectral radiant intensities of tungsten-filament lamps, 400–700 mμ, L. E. Barbrow, J. Opt. Soc. Am., Letter to Editor **49,** No. 11, 1122 (1959).

Paramagnetic resonance spectra of active species. Blue material from hydrazoic acid, W. B. Gager and F. O. Rice, J. Chem. Phys. **31,** No. 2, 564 (1959).

On prediction of system behavior, J. R. Rosenblatt, Proc. of the New York Univ.-Ind. Conf. on Reliability Theory, p. 39 (1958).

Ceramic wafer tubes for modular units, C. P. Marsden, Section on Components and Materials, Electronics **32,** No. 39, 94 (1959).

Magnified and squared VSWR responses for microwave reflection coefficient measurements, R. W. Beatty, IRE Trans. on Microw. Theory and Tech. **MTT–7,** No. 3, 346 (1959).

Argon deposition on a 4.2° K surface, S. N. Foner, F. A. Mauer, and L. H. Bolz, J. Chem. Phys. **31,** No. 2, 546 (1959).

Low-temperature tensile properties of copper and four bronzes, R. M. McClintock, D. A. Van Gundy, and R. H. Kropschot, ASTM Bull. No. 240, 34 (TP177) (1959.)

Kinetic equation for a plasma with unsteady correlations, C. M. Tchen, Phys. Rev. **114,** 394 (1959).

A nomenclature for conformations of pyranoid sugars and derivatives, H. S. Isbell and R. S. Tipson, Science **130,** No. 3378, 793 (1959).

On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. III (General Rayleigh quotient and characteristic roots with linear elementary divisors) A. M. Ostrowski, Arch. Rat. Mech. Anal. **3,** 325 (1959).

On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. IV (Generalized Rayleigh quotient for nonlinear elementary divisors) A. M. Ostrowski, IV Arch. Rat. Mech. Anal. **3,** 341 (1959).

**Publications for which a price is indicated (except for Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (Order by PB number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

A Theory of Radar Scattering by the Moon

T. B. A. Senior and K. M. Siegel[1]

(December 4, 1959)

A theory is described in which the moon is regarded as a "quasi-smooth" scatterer at radar frequencies. A scattered pulse is then composed of a number of individual returns each of which is provided by a single scattering area. In this manner it is possible to account for all the major features of the pulse, and the evidence in favor of the theory is presented. From a study of the measured power received at different frequencies, it is shown that the scattering area nearest to the earth is the source of a specular return, and it is then possible to obtain information about the material of which the area is composed. The electromagnetic constants are derived and their significance discussed.

1. Introduction

Little more than a decade has now elapsed since radar contact with the moon was first established. The first reported contact was by the U.S. Signal Corps in 1946, and this was followed closely by accounts of similar success in Russia. Since then a considerable amount of effort has been expended in attempting to determine the scattering properties of the moon at radar frequencies.

The earlier experiments were limited by the equipment which was available at the time and, in consequence, many of the conclusions reached were largely qualitative. In recent years, however, the use of high-power transmitters in combination with antennas designed specifically for lunar work has provided results the accuracy of which is unquestionable, and these have shown that the scattering properties of the moon cannot be explained in terms of a "rough" scattering theory. The increased transmitted power and the improved receiver sensitivity have also permitted the use of shorter pulse lengths, and with pulses of only a few microseconds in length it has become practicable to distinguish between the returns from different parts of the lunar surface. This in turn has lead to an accurate determination of the radar distance to the moon, and the suggestion has been made that the distance should be used to obtain a more precise value for the earth's equatorial radius.

In many of these experiments a motivating force has been the possibility of employing the moon as a link in a communication system between points on the earth. The successful transmission of cw code was first reported by Sulzer et al. [1][2] in 1952, and subsequently by Trexler [2] who also transmitted voice. It is now well established that this type of link is a practicable proposition and signals can be communicated without appreciable loss of intelligence.

2. Scattering Theories

Optical experiments over many decades have given a fairly comprehensive picture of the major geographical features of the moon, and the general impression received is that the moon's surface is highly irregular. It is therefore natural to think of it as a rough scatterer, and since the intensity of illumination across the disk appears effectively constant, the moon has been characterized by astronomers as a "rough surface" at optical wavelengths. By this it is implied that the scattering properties of the moon can be explained in terms of a surface having a statistically uniform roughness whose scale is the order of a wavelength.

When radar reflections were first observed, it seemed reasonable to expect that these results could also be explained by a rough scattering theory. Thus, in discussing the U.S. Signal Corps experiments, DeWitt and Stodola [3] assumed without question that the moon could be treated as rough, and their analysis of the shape and intensity of the echo pulses was then based on a uniformly illuminated (and reflecting) disk. Similarly, in 1951 Kerr and Shain [4], using pulses of a millisecond or longer, concluded from their returns that "the moon is a rough reflector at this frequency (20 Mc) and hence presumably at all higher frequencies." The same conclusion was also reached by Browne et al. [5] in 1956 from a study of the echo at 120 Mc with 30-millisecond (msec) pulses, and it was not until shorter pulse lengths came into use that the validity of rough scattering theories was seriously questioned. It was then found that most of the return at radar wavelengths comes from only a small portion of the lunar surface, and in a summary paper published in 1957 Lovell [6] stated that "recent experiments have shown the 'uniformly bright' conclusion to be erroneous."

From an examination of pulse shapes such as those obtained by Yaplee et al. [7], the present authors felt that a scattering process similar to that for a smooth body would explain the dominant features of the returns. Thus, in 1958, a scattering

[1] Radiation Laboratory, Department of Electrical Engineering, University of Michigan, Ann Arbor, Mich.
[2] Figures in brackets indicate the literature references on page 228.

536182—60——1

theory was proposed [8] in which the starting point was the idea of a body which is essentially smooth as regards its radar scattering properties, and to take into account the existence of several scattering areas, the term "quasi-smooth" was applied to the lunar surface. In this paper we shall discuss the theory at some length and, in so doing, attempt to clarify those steps which, to judge from criticisms received, would appear to have caused difficulty. The fact that more experimental data has become available since the theory was originally propounded also enables us to place some of the conclusions on a somewhat firmer basis, while amending others.

3. Fading Phenomena

In radar returns from the moon, two distinct types of fading are apparent. The more rapid fading has a period of only a few seconds and was first observed by DeWitt and Stodola [3], and later by Kerr and Shain [4] who attributed it to the moon's libration. This has since been confirmed by Evans et al. [9] from a consideration of the diurnal variation in the fading rate. The second type of fading has a much longer period (as much as 20 to 30 min) and is most clearly seen in results of experiments at the lower frequencies. It has been shown by Murray and Hargreaves [10] that these fades are consistent with the Faraday rotation of the plane of polarization produced by the earth's ionosphere. This explanation is now well accepted, but in view of the importance of the longer period fades in connection with the scattering process at the moon's surface, it is necessary to consider them in rather more detail.

3.1. Faraday Rotation

When a linearly polarized electromagnetic wave passes through an ionized medium in the presence of a magnetic field its plane of polarization is rotated. For a linearly polarized signal which is directed towards the moon, the polarization is first rotated on passage through the ionosphere and then, after reflection at the moon with an appropriate phase change, the polarization is again rotated on returning through the ionosphere. As the ionosphere changes, so will the rotation, and this in turn will produce a variation in the level of the signal received with an antenna of fixed linear polarization. Given a sufficient period of observation, the signal will pass through a maximum and this will occur at the same instant of time as would a minimum in an antenna of orthogonal polarization. In the extreme circumstance in which no depolarization took place at the moon's surface, this minimum would be a true zero and then, in theory at least, all the returned energy would be received by the first antenna.

The time between successive maximums or minimums is the Faraday rotation period and has been measured by many experimenters. Apart from some ambiguities it leads directly to a value for the integrated electron density along the propagation path and has been used for this purpose by, for example, Browne et al. [5].

3.2. Fading Minimums

The above facts are relevant to the analysis of data such as that obtained by Blevis and Chapman [11]. Using a cw system operating at a frequency of 488 Mc, the signals scattered back by the moon were received with two orthogonally polarized antennas, and samples of their data are displayed[3] in figure 1. It will be observed that when one polarization is at a maximum, the orthogonal polarization is at a minimum, and we have seen that this is a natural consequence of the Faraday rotation. More remarkable, however, is the ratio of the maximum to minimum signal levels, which is found to be approximately 16 db [12]. A ratio as large as this is not compatible with scattering from a rough surface. Any rough surface serves to depolarize the incident field, and for a surface on which bumps are placed whose magnitude is of order one-half wavelength, the depolarization effect will be almost complete; the returns received with two orthogonally polarized antennas will then differ by only a few decibels independently of the polarization of the transmitted field. In view of the large ratio found in the case of the moon, it can now be concluded that as regards the major sources of the scattered energy, the scattering is akin to that from a smooth surface; and in particular, the surface cannot possess a large number of (contributing) irregularities whose magnitude is comparable with the wavelength.

Although these conclusions have been reached by a study of the returns in two orthogonally polarized antennas, the same argument can be applied to the signal received on one linear polarization alone. If the maximum signal level is compared with the minimum (received at a later time), and if the interval of time is the Faraday rotation period, the ratio is again a measure of the roughness. In this respect at least, the ionosphere has a desirable effect in enabling us to dispense with one of the receiving antennas.

3.3. An Experiment

To emphasize the argument outlined above, an experiment was carried out in the laboratory using a sphere possessing bumps. The diameter of the sphere was 10λ and the size of the bumps was of the order of λ. A picture of the sphere is shown in figure 2 and some of the results are given in figure 3. These can be summarized as follows:

Relative scattering cross section in decibels

	σ_{VV}	σ_{VH}
Smooth sphere	0	−30
"Perturbed" sphere	−7	−13

The cross sections for the perturbed sphere were averaged over 60° in aspect.

---

[3] The authors thank Dr. John Chapman of the Communications Laboratory, Defence Research Telecommunications Establishment, for allowing figures 1 and 4 to be copied from presently unpublished data of Blevis and himself.

NOTE ADDED IN PROOF: Since this paper was written the authors have learned that Chapman and Blevis have submitted for publication in this journal a paper containing the data.

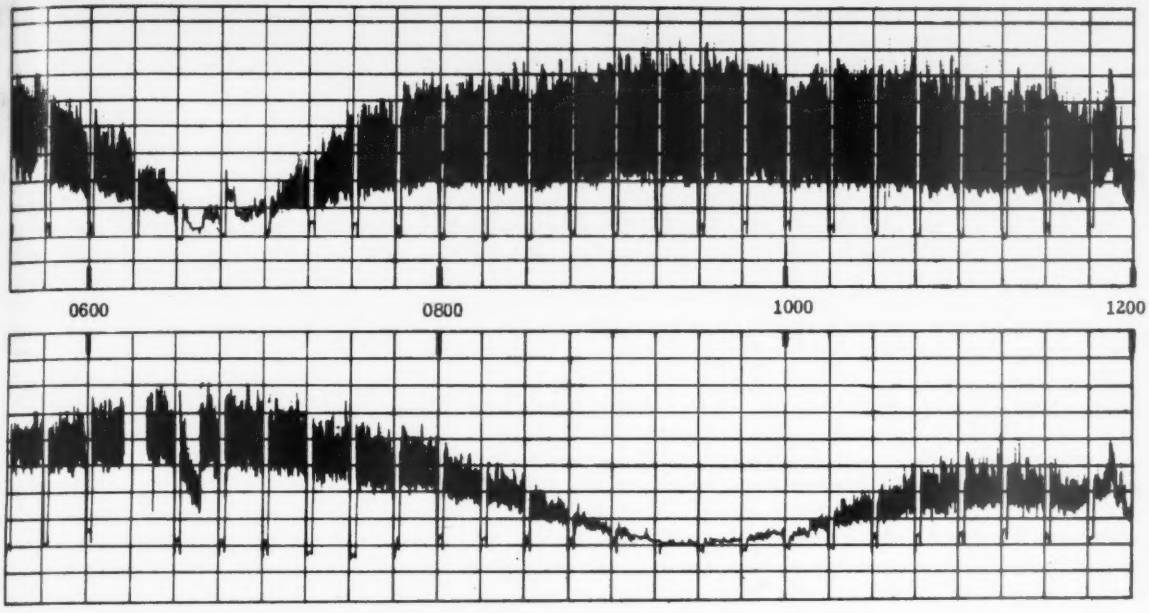


FIGURE 1. *Faraday rotation record, September 15, 1959.*

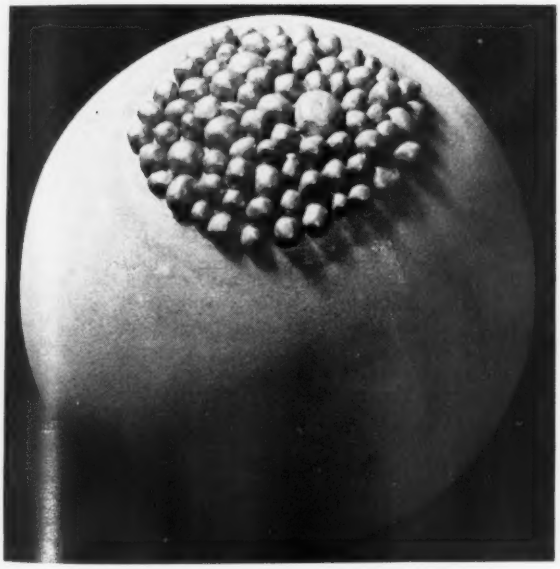


FIGURE 2. *Sphere with bumps.*

It is seen that for a sphere without any bumps, the direct return (transmitter and receiver polarizations the same) exceeds the cross-polarized return by 30 db; whereas when the bumps are present, the cross-polarized return is increased by 17 db, and the direct return decreased by 7 db, so that when averaged over a range of 60° in aspect, the difference be-

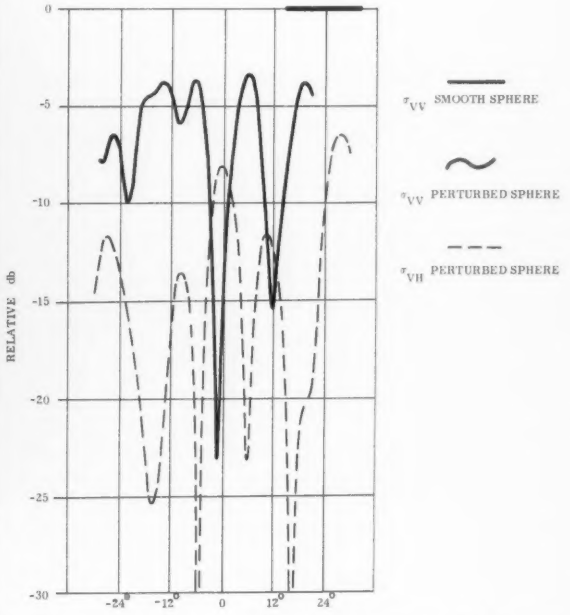


FIGURE 3. *Effects of bumps on the backscattering cross section of a 10λ sphere.*

tween the polarizations is no more than 6 db. The measured difference decreases with decreasing size of bumps until the bumps become small compared with the wavelength. Accordingly, if the moon were a sphere covered with wavelength bumps, the ratio of the maximum and minimum signals in the Faraday rotation period would not exceed (about) 6 db. This is in marked contrast to the ratio which is actually observed.

3.4. Double Fades

The question of double fades is not directly connected with the development of our theory, but it may be of interest to discuss the matter here.

At the XIIth General Assembly of URSI at Boulder, Colorado (September 1957), Chapman pointed out that in his moon returns he sometimes obtained minimums occurring simultaneously in two orthogonally polarized receiving antennas. These phenomena were termed "double fades" and can be observed in the sample records shown in figure 4. They were discussed in some detail at the Assembly, and both there and in later meetings were felt to violate the conservation of energy. This, however, is not so, and indeed double fades can be reproduced in the laboratory.

To explain the origin of these fades, we first employ a statement of the conservation of energy. For a fixed aspect of target, for a given transmitter polarization α, and for all receiver polarizations β, the equation

$$\sigma(\alpha,\beta)+\sigma\left(\alpha,\beta+\frac{\pi}{2}\right)=\text{constant} \qquad (1)$$

holds, where $\sigma(x,y)$ is the scattering cross section when the transmitting antenna is polarized in the direction x and the receiving antenna is polarized in the direction y. Equation (1) is a direct consequence of conservation of energy and is an identity for all β (see app.). It should be noted that as a function of aspect (which can be varied by rotating the body with the transmitting and receiving antennas fixed), this says nothing about whether two minimums can or cannot occur simultaneously in two orthogonally polarized receivers. Indeed, it is to be expected that for a nonsymmetrical body there will be a particular aspect such that the total backscattered energy is a minimum, and the constant in eq (1) is then a function of aspect. If at this minimum aspect we invoke the theorem represented by eq (1), we merely find that the total power received by two orthogonal antennas is independent of their polarization. Both $\sigma(\alpha,\beta)$ and $\sigma(\alpha,\beta+(\pi/2))$ will then be small for any chosen polarization β, and both may well be smaller than they are for neighboring target aspects. In this way a double fade could occur.[4] It in no way violates eq (1), nor the conservation of energy, and is not an unreasonable thing to occur with a target of asymmetrical shape.

4. Pulse Considerations

4.1. Short Pulses

Further evidence in favor of a quasi-smooth moon theory is provided by the shape of short pulse returns. When the transmitted pulse is only a few microseconds in length, the return appears to be dominated by a small number of spikes. The onset of the return is relatively sharp and the time of its occurrence is consistent with reflection from that portion of the moon which is nearest to the earth [13]. In many cases at least the spike which is associated with the leading edge of the pulse represents the peak power return.

In a series of experiments using 2-μsec pulses, Yaplee et al. [7] found that each of these spikes had a width of from 2 to 4 μsec and were separated in time by 2 to 4 μsec. The close agreement between the width of the spikes and the transmitted pulse length suggests that the source of each spike is a single scattering area. If this is so, the spikes should possess a high pulse-to-pulse correlation and a narrow power spectrum. These have been determined by Yaplee et al. [14], and by measuring the amplitude of the return as a function of distance from the leading edge it has been shown that all the first few spikes have a high correlation. The correlation decreases with increasing distance from the leading edge and, ultimately, the amplitudes assume a random distribution. As a further check the power spectra were also obtained and, as to be expected, were relatively narrow for the first few spikes.

[4] Since this explanation was proposed we have learned that Dr. Chapman has been thinking along similar lines.

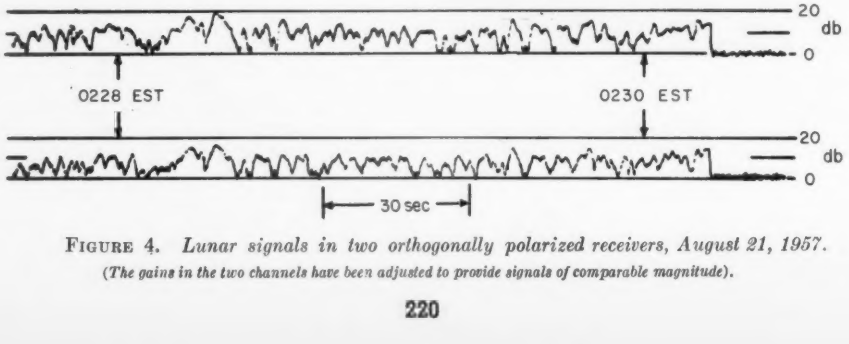


FIGURE 4. *Lunar signals in two orthogonally polarized receivers, August 21, 1957.*
(The gains in the two channels have been adjusted to provide signals of comparable magnitude).

The above results cannot be explained by a rough scattering theory, nor by one in which many scattering areas contribute towards every individual spike. On the other hand, a theory in which each leading spike is produced by a single scattering area is entirely in accordance with all the available data. The picture which now emerges has a small number of scattering areas distributed over the moon and these alone are responsible for the major features of the return. The fact that the pulse-to-pulse correlation decreases with increasing distance from the leading edge of the return is then a natural consequence of the increasing area of illumination with depth on the moon, and hence, of the increased probability of two scattering areas contributing simultaneously.

If the scattering areas were distributed over the entire moon's surface, the return from even a short pulse could persist for as much as 11.6 msec, with the last contribution coming from the areas on the limbs. In the early moon experiments the long pulse lengths and the low transmitted powers made difficult an estimate of the amount of elongation, but Kerr and Shain [4] reported an observed elongation with even millisecond pulses. In recent years, however, the use of microsecond pulses has made the actual elongation apparent, and it is now accepted that the bulk of the return is produced by only a small portion of the moon. Thus, Trexler [2] has observed that 50 percent of the power comes from the first 5 miles of depth or a circle only 105 miles in radius. This is almost exactly one tenth the radius of the moon. The most distant echoes seen were only 100 miles back. Similarly, Lovell [6] states that the main scattering takes place within an area whose radius is about one third that of the moon.

It follows that the scattering areas are not uniformly distributed over the moon's surface but are concentrated near to the center. The return from portions of the moon far from its center is relatively small and, for example, at times subsequent to the onset of the return corresponding to reflection from the limbs, the signal (in the rare cases when it has been seen) is down by 30 to 35 db from that produced by the main scattering areas. In fact, it is only with radars such as those at Millstone Hill and Jodrell Bank that the dynamic range has been sufficient for the limb return to be detected.

4.2. Long Pulses

If the length of the transmitted pulse is increased, the character of the return is changed. Instead of a few major spikes superimposed upon a rough but decreasing background, there now exists a progressive buildup in the return. Given the existence of scattering areas it is to be expected that as the pulse length increases the maximum level of the return will also increase due to the simultaneous illumination of more and more areas. On the other hand, because of the concentration of these areas near the center of the moon, the increase cannot continue indefinitely, and ultimately a pulse length is reached at which the power return is the same as that obtained with cw. This behavior is indeed found. With a long transmitted pulse the leading edge of the return is still sharp and again corresponds to scattering from that portion of the moon which is nearest to the earth, but while this area is still being illuminated the return from the next scattering area comes in and interferes with the first. As time progresses more returns appear, and these will interfere to produce a variety of peaks which are superimposed upon an increasing background level. This will continue for a time equal to the transmitted pulse length, or until the cw return has been achieved, whichever is the sooner. With increase of time beyond the transmitted pulse length, the level of the return decreases and disappears into noise after (about) 200 μsec. In a return of this type all the peaks other than the first must be attributed to interference.

As stated above, an increase in the pulse length is usually accompanied by an increase in the peak power return. Conversely, a decrease in the pulse length is generally reflected by a decrease in the return, and this leads to the concept of modulation loss. The modulation loss was first introduced by Trexler [2] and is essentially a measure of the reduction in the peak power return when using finite pulse lengths, as compared with the level of the cw return. Based upon many experiments at different frequencies using a variety of pulse lengths, Trexler has determined values for the loss. These have been communicated by Youmans [15] and show a modulation loss that increases from 1 db for a pulse length of 200 μsec, through 8.5 db at 30 μsec, and 17 db at 10 μsec, to 22 ± 2 db at 2 μsec. The number of points is not enough to describe the precise variation at small pulse lengths, but it would appear that the loss is effectively constant for pulse lengths less than approximately 5 μsec. With pulses of this duration, the contribution from the first scattering area is isolated from the rest. In contrast, pulses of 300 μsec are sufficient to realize the peak cw return.

4.3. Scattering Areas

In order to estimate the number of major scattering areas two different methods are available. The first of these consists of merely counting the larger peaks in a short pulse return. With this approach the number obtained lies somewhere between 20 and 30, but inaccuracies do arise because of interference between signals from scattering areas in the same annular regions of the moon. In other words, a one-to-one correspondence between peaks and scattering areas may not exist for all portions of a returned pulse.

The second method is based upon the modulation loss and involves a detailed examination of the power levels used. To begin with we remark that the measured values for the modulation loss were found by restricting the times of observation to ones at which an appreciable signal is received. Such a statement has many ramifications and not only implies good propagation conditions, but a straight average over time. Thus, the cw level which is employed as a reference is an average over time for those

periods at which a large return appears. This level exceeds the one corresponding to random excitation of the various scattering areas.

As regards the pulse returns, the power level used is the average over pulses of the maximum for each. With very short pulses (2 to 5 μsec, say), this average is merely the return from the first scattering area. As the pulse length increases more areas are illuminated, but with short pulses at least the average peak level is almost that appropriate to in-phase addition of the individual returns. Naturally, the amount by which it falls below the in-phase level increases with increasing pulse length, but because the scattering areas are not large in number, this difference never becomes more than (about) 7 db. In consequence, the average peak level remains considerably above the level corresponding to random excitation. The figure 7 db was obtained from results at 201 and 915 Mc using millisecond pulses [16] and represents the difference between the largest signal observed and the mean level in the sense described above. It is natural to imagine that this large return occurred when the individual signals were precisely in-phase, and the assumption that this was so provides an estimate for the number of major scattering areas. For this purpose we assume that each scattering area gives rise to a signal of the same magnitude and, in addition, that the peak return from a 2-μsec pulse comes from a single area. All the areas together then produce a peak (or in-phase) return which is $22+7=29$ db larger than that from one, indicating that there are 28 areas in toto. The result is in good agreement with that obtained by direct counting of spikes, and this is in spite of the inaccuracies inherent in both methods. A more detailed discussion of the subject is out of place here, but the above arguments are the basis for our present belief that the number of major scattering areas is between 20 and 30.

No statement has yet been made about the physical nature of the scattering areas. Although much work remains to be done in this connection, some information can be obtained by an extension of the above analysis. If the modulation loss curve is modified in such a way as to give results appropriate to in-phase addition, it is possible to estimate the variation in the total number of scattering areas in an annular region for which the slope of the mean lunar surface changes by a fixed amount (for example, 1°) from the tangent plane at the nose. It is found that the number of areas per degree of slope decreases uniformly with increase of slope beyond about 2°. This occurs in spite of the increasing area of these regions and suggests that whereas the specular areas nearest the center of the moon are portions of spherical surfaces, the areas further back are possibly craterlike. As the depth increases the decreasing probability of finding surfaces suitably alined is the reason why the level of a short pulse return falls below its peak by 30 or more decibels after 600 μsec (approx) from the onset of the return. In effect, therefore, all the scattering areas are located within 300 μsec of the center of the moon, corresponding to a depth of 30 miles.

4.4. Pulse Elongation

A direct consequence of the existence of many scattering areas is the elongation of any pulse returned from the moon. We have seen that this can be as much as 600 μsec and, indeed, if a short pulse return is examined, the general impression received is of a few major spikes superimposed on a rough but decreasing background. This background will be called a "tail" and its general features are the same for pulses of any length. The peaks have already been discussed in some detail, and we shall now consider possible mechanisms which could produce other pulse distortion and, in particular, generate a tail.

It is first remarked that for a smooth sphere of radius a the return from an incident pulse has a tail which is a function of both the pulse length and the wavelength. If, for example, the pulse length is of order a and, in addition, $2\pi a/\lambda$ has a value near to unity, the analysis given in [19] shows that a tail will persist with appreciable amplitude for a considerable length of time. For a body which has many scattering areas it is to be expected that the relevant dimension is the size of the scattering areas, and if this is now comparable with the wavelength, each individual return may well possess a tail. The precise amount of tail will depend upon the type of scattering object, but since the dimensions of the major scattering areas on the surface of the moon are almost certainly measured in tens (or hundreds) of wave-lengths, this effect is unlikely to be significant.

In practice, there are many other effects which are almost certainly more important. Any roughness of the lunar surface will produce a residual signal and, in addition, pulse degradation will be caused by differential action on the frequency components by the propagation medium (for example, the Faraday rotation is frequency dependent). Finally, there is the limited receiver bandwidth to be considered. In most of the experiments the bandwidth has been less than the optimum for the pulse length, and, mathematically at least, this could be the source of a large amount of the tail. Taking, for example, the experiments carried out by Yaplee et al. [7] using 2 μsec pulses, a straightforward calculation shows that the bandwidth of 350 kc could generate a tail which only decays to 6 percent of the peak after 16.7 μsec. Under these circumstances, a detailed discussion of the tail contributions is somewhat academic.

5. The Measured Power Return

In the last few years many measurements have been made of the transmission loss in radar scattering by the moon, and from these results the scattering cross section can be obtained. The quantity of data is now increasing rapidly, and the results which are available at this time are listed in table 1. The cross sections are there expressed as fractions of πa^2 with $a=1.74\times10^6$ m. It will be recognized that πa^2 is the cross section of a perfectly conducting sphere

TABLE 1

	Wavelength	$\sigma/\pi a^2$	Probable error (where known)
	m		*db*
(i)	0.10	4×10^{-4}	
(ii)	.14	3×10^{-4}	±4 or 5
(iii)	.33	9×10^{-2}	±3
(iv)	.61	3×10^{-2}	±3
(v)	.73	7×10^{-2}	±3
(vi)	.75	1×10^{-1}	±3
(vii)	1.00	$(5\sim9)\times10^{-2}$	±4
(viii)	1.49	7×10^{-2}	±3
(ix)	1.50	$(6\sim10)\times10^{-2}$	±4
(x)	1.99	5×10^{-2}	
(xi)	2.5	1×10^{-1}	±3
(xii)	3.0	1×10^{-1}	±3

whose radius is that of the moon. In interpreting the data an essential factor is the pulse length. Whereas the measurements at 3,000 Mc were carried out with pulse lengths of 5 and 2 μsec, all the others were based on an analysis of cw (or very long pulse) returns, and this accounts for the significantly lower values in (i) and (ii).

Of the above values of σ, the first was obtained by Hey and Hughes [17] of the Royal Radar Establishment at a frequency of 3,000 Mc using 5-μsec pulses. The second is due to Yaplee et al. [7] of the U.S. Naval Research Laboratory, and is for 2,860 Mc with 2-μsec pulses. The third and eighth were found by Aarons et al. [16] of the Air Force Cambridge Research Center using millisecond pulses at frequencies of 915 and 201 Mc, respectively. The fourth measurement[5] was made by Blevis and Chapman [11] of the Defence Research Telecommunications Establishment at a frequency of 488 Mc using a cw system, and the fifth is also a cw result obtained by Fricker et al. [18] of Lincoln Laboratory with a frequency of 413 Mc.

The sixth and twelfth values were found by Leadabrand [19] of the Stanford Research Institute using millisecond pulses at frequencies of 400 and 100 Mc, respectively. The seventh is due to Trexler [2] of the U.S. Naval Research Laboratory and has been deduced from his statement that the cw transmission loss is 258 db when calculated on the basis of isotropic transmitting and receiving antennas. The two values given for σ in (vii) and (ix) are for a moon at perigee and apogee, respectively. The ninth result is the average obtained from many cw measurements made at or near a frequency of 200 Mc, and has been stated by Trexler [2] as a transmission loss of 254 db. The tenth measurement was made by Webb [20] of the University of Illinois at a frequency of 151 Mc using a cw system. The eleventh result was obtained by Evans [21] of Jodrell Bank with a pulse length of 30 msec at a frequency of 120 Mc, and is a refinement of that originally measured by Browne et al. [5].

In addition to the above there is a "summary measurement" which indicates the general dependence of the cross section on wavelength. From a review of the data available to him, Trexler [2] states that for frequencies between 20 and 3,000 Mc the average cw transmission loss increases at a rate of 6 db per octave, passing through the value 258 db at 300 Mc. Reference to the radar equation shows that this implies a cross section which is wavelength independent.

A meaningful comparison of all the available data can only be carried out if the cross sections can be reduced to those which are appropriate to a common pulse length. It will be observed that most of the values for σ were obtained from either cw or long pulse experiments, and as such they represent the power contributed by many scattering areas. To discuss the properties of a single scattering area, it is necessary to know the power associated with one area alone, and this is only the peak return in the case of very short pulses. It would clearly be desirable to confine the analysis to data which was actually obtained with pulses of this length, but since the number of these experiments is still very small, we are forced to make use of all available data. To do this, "correction factors" are applied to the scattering cross sections to reduce them to the values for microsecond pulses. Using Trexler's figure for the modulation loss associated with a pulse of duration 5 μsec or less, the modified data is then as follows:

(i) $\lambda=0.10$ m, $\sigma=4\times10^{-4}\ \pi a^2$
(ii) $\lambda=.14$ m, $\sigma=3\times10^{-4}\ \pi a^2$
(iii) $\lambda=.33$ m, $\sigma=5.4\times10^{-4}\ \pi a^2$
(iv) $\lambda=.61$ m, $\sigma=1.9\times10^{-4}\ \pi a^2$
(v) $\lambda=.73$ m, $\sigma=4.3\times10^{-4}\ \pi a^2$
(vi) $\lambda=.75$ m, $\sigma=6.3\times10^{-4}\ \pi a^2$
(vii) $\lambda=1.00$ m, $\sigma=4.7\times10^{-4}\ \pi a^2$
(viii) $\lambda=1.49$ m, $\sigma=4.2\times10^{-4}\ \pi a^2$
(ix) $\lambda=1.50$ m, $\sigma=5.3\times10^{-4}\ \pi a^2$
(x) $\lambda=1.99$ m, $\sigma=3.0\times10^{-4}\ \pi a^2$
(xi) $\lambda=2.5$ m, $\sigma=6.3\times10^{-4}\ \pi a^2$
(xii) $\lambda=3.0$ m, $\sigma=6.3\times10^{-4}\ \pi a^2$

These results are interpreted as the scattering cross section of the first scattering area, and are plotted in figure 5. Unfortunately it is difficult to estimate their accuracy. Even with the original data the probable error is not always known, but where it is it has been noted in table 1. It is seen that a typical error is ±3 db. In addition to these inaccuracies there are those which are inherent in trying to deduce the results of a short pulse experiment from a long pulse or cw measurement. In the absence of more definite information, we have therefore associated an uncertainty of ±3 db with all the modified data, and have denoted these limits by arrows in figure 5. It would seem likely that this is an underestimate, but nevertheless the consistency of the results gives added faith in both the original data and in Trexler's measurement of the modulation loss.

6. Theoretical Considerations

To obtain a theoretical expression for the cross section of the first scattering area, it is necessary to know the physical shape of the scatterer and also the material of which it is composed. In the original exposition of the quasi-smooth moon theory [8], it was

[5] NOTE ADDED IN PROOF: We have been informed by Dr. Chapman that a more detailed study of their data has increased their value for the cross section of the moon to $5\times10^{-2}\ \pi a^2$. This has the effect of making still better the agreement between theory and experiment (see fig. 5).

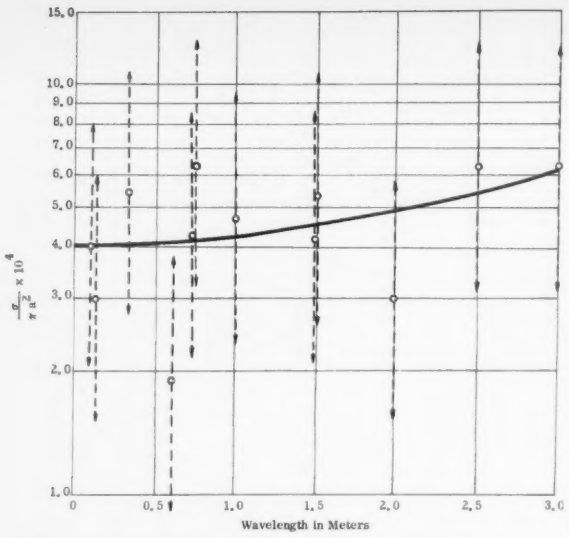


FIGURE 5. *Comparison between theory and experiment.*

assumed that the return from this area is a specular contribution, but the suggestion has been made that facets, or even conelike projections, should also be considered as sources of return. To allow for this possibility, a formula for the cross section will now be derived of sufficient generality to embrace several types of scatterer.

6.1. Basic Formulas

Let us consider first a perfectly conducting body of arbitrary convex shape on which a plane wave is incident. It will be assumed that the receiver is in the far field of the body regarded as a scatterer.

Since the body is perfectly conducting, the reflection coefficient is unity, and the leading term in the asymptotic expansion of the cross section for small wavelengths is given by geometrical optics. If the body has no corners or facets perpendicular to Poynting's vector, the geometrical optics cross section is wavelength independent and is simply

$$\sigma=\pi\, R_1\, R_2, \tag{2}$$

where R_1 and R_2 are the principal radii of curvature of the specular region. For backscattering, the center of the specular region is that point on the body which is nearest to the receiver. In the particular case in which $R_1=R_2=a$, eq (2) reduces to

$$\sigma=\pi a^2, \tag{3}$$

which will be recognized as the cross section of a perfectly conducting sphere of radius a.

If the body is not smooth and convex, but is still perfectly conducting, other portions may contribute towards the scattering cross section, and while small scale irregularities will not be considered, any large scale perturbation must be taken into account. Falling into this category is a corner reflector or large flat area, the cross section of which is

$$\sigma=\frac{A}{\lambda^2}, \tag{4}$$

where A is proportional to the square of the area. For positions of the receiver in the direction of peak return, A is essentially wavelength independent. The equivalent structure on the surface of the moon could be a smooth mountain face taken either by itself (to give a flat area) or in combination with the level lunar surface (to produce a corner reflector). Because of the narrow beamwidth, it seems unlikely that such a scatterer could provide a significant contribution to the actual moon's return over any range of aspects, but in the interests of generality the possibility will not be excluded.

A second type of structure is a conelike projection of small included angle (the lunar equivalent might be a mountain peak), and the cross section of this is

$$\sigma=C\lambda^2, \tag{5}$$

where C is independent of the wavelength. If both types of scatterer are present on the body, the total cross section can be written in the form

$$\sigma=\frac{A}{\lambda^2}+B+C\lambda^2, \tag{6}$$

with A, B, $C\gg 0$. In this expression, B represents the geometrical optics cross section of the unperturbed body.

It will be appreciated that in writing down eq (6) we have neglected interference between the individual scatterers. In its application to the moon, the equation is only required in the analysis of the leading part of the pulse, and to contribute towards this part all the relevant scatterers must lie in the same general area. The interference which then exists will show up as a fluctuation in the power level of individual pulses as a function of aspect. Since no fluctuation appears which can be attributed to this effect, it seems likely that one type of scatterer and, moreover, only one scatterer of that type, is responsible for the major contribution towards the leading part of the pulse. Accordingly, if the cross section in eq (6) is compared with the experimental data, we should expect to find that one of A, B, and C dominates the other two.

Unfortunately, eq (6) is not immediately applicable to the analysis of the data. In the first place, the moon is not perfectly conducting nor is it likely to be homogeneous. In addition, many of the experiments have been carried out using beamwidths which do not always illuminate the whole moon and, finally, there is the near-zone effect. As pointed out in reference [8] the distance between the earth and the moon is substantially less than $4\pi a^2/\lambda$, where a is the moon's radius, so that any receiver on the earth

will necessarily be within the near zone of the moon regarded as a scatterer. The effect, if any, of these conditions upon the scattering cross section will now be considered.

6.2. Near-Zone Effect

After it was appreciated that the earth is in the near zone of the moon at all frequencies which will penetrate the ionosphere, Dr. V. H. Weston of this laboratory obtained the leading term in the small wavelength expansion for the cross section of a perfectly conducting sphere with transmitter and receiver located arbitrarily.

Taking the transmitter to be a point source at a distance r_1 from the nearest point on the surface of the sphere and the receiver at a distance r_2 from the same point and in the same direction, Weston's analysis [22] yields a cross section

$$\sigma=4\pi\left(\frac{ar_1r_2}{a(r_1+r_2)+2r_1r_2}\right)^2 \tag{7}$$

where a is the radius of the sphere. There are now two distinct cases that can be treated. If the transmitter and the receiver are at the same position so that $r_1=r_2(=r$, say), eq (7) becomes

$$\sigma=\pi\left(\frac{ar}{a+r}\right)^2. \tag{8}$$

When $a>>r$,

$$\sigma=\pi r^2,$$

which is therefore the cross section of a sphere for a transmitter and receiver near to the surface and also the cross section of an infinite flat plate for a transmitter and receiver anywhere, and when $a<<r$, eq (8) gives

$$\sigma=\pi a^2.$$

A second set of results can be deduced from eq (7) when the incident field is a plane wave. By taking the limit of (7) as $r_1\to\infty$ (it is sufficient that $r_1>>a, r_2$), we have

$$\sigma=4\pi\left(\frac{ar_2}{a+2r_2}\right)^2 \tag{9}$$

and for $a>>r_2$

$$\sigma=4\pi r_2^2,$$

which differs by a factor 4 from the corresponding result for spherical wave illumination. Of course, when $r_2>>a$, eq (9) reduces to

$$\sigma=\pi a^2$$

in agreement with eq (3).

The above analysis shows that for cw illumination the leading term in the expansion for the cross section of a sphere is πa^2 for both plane and spherical waves, and is independent of the position of the receiver providing its distance from the sphere is large compared with the radius. In reference [22] it is also shown that the same formula gives the scattering cross section which would be obtained with pulses of any length.

These results have an obvious interpretation in terms of the cross section of the specular area, and in order to realize the full cross section it is only necessary for the incident field to cover this particular area. In consequence, the beam width can be appreciably less than the angle subtended by the sphere without affecting the peak return. Moreover, because the moon's return is given by a finite number of scattering areas, we need only illuminate these to obtain the total power return, and since the areas are confined to a region about 500 miles in extent, a beamwidth of approximately 8 min of arc should be sufficient to realize all the available power.

6.3. Reflection Coefficient

The formulas which have been obtained so far only apply to perfectly conducting bodies, and we shall now consider the effect of taking into account the material constants of the body.

For a body which is not perfectly conducting but is composed of a homogeneous isotropic material, the leading term in the small wavelength expansion differs from the optics cross section by a factor which depends upon the electromagnetic constants of the surface and is, in fact, the power reflection coefficient $|R|^2$. In the case of backscattering from a flat plate or specular area, the maximum return occurs at normal incidence, and since the dimensions of the scatterer are large compared with the wavelength, the reflection coefficient is the same as for an infinite slab of the material.

If a plane electromagnetic wave is incident normally on such a slab,

$$R=\frac{1-\sqrt{\frac{\mu_0}{\mu}\left(\frac{\epsilon}{\epsilon_0}+i\frac{s}{\omega\epsilon_0}\right)}}{1+\sqrt{\frac{\mu_0}{\mu}\left(\frac{\epsilon}{\epsilon_0}+i\frac{s}{\omega\epsilon_0}\right)}} \tag{10}$$

(see, for example, reference [23]), where ϵ, μ, and s are the permittivity, permeability, and conductivity, respectively, of the material. The subscript "o" denotes the same quantities for free space and ω is the frequency. Rationalized mks units are employed with a time dependence $\exp(-i\omega t)$.

If the material is primarily dielectric in behavior at the relevant frequencies (as seems probable for the moon), then

$$s<<\omega\epsilon$$

and

$$R\simeq\frac{1-b-ib'}{1+b+ib'},$$

where

$$b=\sqrt{\frac{\mu_0\epsilon}{\mu\epsilon_0}}\quad\text{and}\quad b'=\frac{s}{2\omega\epsilon_0}\sqrt{\frac{\mu_0\epsilon_0}{\mu\epsilon}},$$

giving

$$|R|^2\simeq\frac{(1-b)^2+b'^2}{(1+b)^2+b'^2}.$$

In the denominator the second term is certainly small compared with the first, and hence

$$|R|^2 \simeq \left(\frac{1-b}{1+b}\right)^2 + \left(\frac{b'}{1+b}\right)^2.$$

Since $\omega = 2\pi c/\lambda$, this equation can be written as

$$|R|^2 \simeq d + \lambda^2 d' \tag{11}$$

with

$$d = \left(\frac{1-b}{1+b}\right)^2 \tag{12}$$

and

$$d' = \left(\frac{b}{1+b}\frac{s}{4\pi c\epsilon}\right)^2. \tag{13}$$

It seems reasonable to assume that d will not be identically zero.

For a flat plate or specular area the cross section for infinite conductivity must be multiplied by the reflection coefficient $|R|^2$, and while the values of d and d' may be different for each individual scatterer, no confusion should arise if distinguishing marks are omitted.

In the case of scattering by a conelike projection, the reflection coefficient cannot be obtained from a consideration of normal incidence on an infinite slab, but from a study of the complete optics solution for a cone of imperfect conductivity, the corresponding expression for $|R|^2$ is found to have the same wavelength dependence as appears in eq (11). Although the coefficients are no longer given by eqs (12) and (13), a detailed derivation is unnecessary in view of the conclusions reached in section 7.1. If the same symbols d and d' are used to denote these coefficients, the combined cross section of the various types of scatterers can be written in the following general form:

$$\sigma = \frac{Ad}{\lambda^2} + (Ad' + Bd) + (Bd' + Cd)\lambda^2 + Cd'\lambda^4. \tag{14}$$

In order to apply this result to the analysis of the data, it is necessary to assume that d and d' are independent of wavelength. For the material of which the moon's surface is composed it seems reasonable to assume that ϵ, μ, and s do not vary rapidly as a function of frequency, and the expression for $|R|^2$ can then be used providing the material of the first specular area can be regarded as homogeneous. In particular, there is no requirement that the whole moon should be effectively homogeneous.

7. Comparison With the Data

The experimental data for the scattering cross section of the first specular area will now be used to determine the most probable values for the constants which appear in eq (14). In this way it is possible not only to determine the dominant type of scatterer which is responsible for the return, but also to obtain information about the material of which the scattering area is composed.

7.1. Least Squares Analysis

Over the range of wavelengths for which the data is available, the variation in the cross section is relatively small, and this suggests that the dominant term in eq (14) is wavelength independent. This is confirmed by a least squares analysis in which the above expression for σ is compared with the measured data. The values obtained for the terms are

$$Ad = -8.94\times10^{-7}\times\pi a^2 \text{ m}^4,$$

$$Ad' + Bd = 4.45\times10^{-4}\times\pi a^2 \text{ m}^2,$$

$$(Bd' + Cd) = -1.03\times10^{-5}\times\pi a^2 \text{ dimensionless},$$

$$(Cd') = 3.71\times10^{-6}\times\pi a^2 \text{ m}^{-2},$$

where $a = 1.74\times10^6$ m. In the range of λ for which we have experimental values we find that the second term exceeds all others and for the smaller wavelength experimental points the second term exceeds all others by at least an order of magnitude. In particular, Ad is both small and negative, and can therefore be neglected. Since $d \neq 0$, it follows that A can be taken zero, and this rules out the possibility of facets making a major contribution towards the measured values of σ. In addition, Cd' is extremely small, and although this does not determine that C itself is negligible, it does enable us to remove a further wavelength dependence from the formula for σ.

By repeating the least squares analysis using only the λ^0 and λ^2 terms, it is found that

$$Bd = 4.04\times10^{-4}\times\pi a^2 \text{ m}^2 \tag{15}$$

$$(Bd' + Cd) = 2.37\times10^{-5}\times\pi a^2 \text{ dimensionless}. \tag{16}$$

The specular area still provides the dominant contribution and since neither Bd' nor Cd can be negative, the effect of any conelike projection is smaller by at least a factor of 17 for $\lambda < 1$ m. On the other hand, no argument of this type can distinguish between the return from such a projection and the conductivity effect of the first specular area. As previously remarked, a variation in amplitude in the leading part of pulses could indicate the existence of more than one scatterer, but the fact that no variation exists only precludes the possibility of two or more scatterers giving *comparable* returns. In the present case we have already seen that the specular area easily outweighs all the other scatterers, and on the ground that a single scattering object provides the most attractive picture, the assumption will be made that C can be neglected. It should be noted that the presence (or absence) of any conelike contribution in no way affects the value of Bd.

The above assumption implies that the leading part of a pulse comes from a single specular area.

and since the scattering cross section is then

$$\sigma=\pi R_1 R_2 |R|^2, \tag{17}$$

where R_1 and R_2 are the radii of curvature, the lack of any appreciable pulse-to-pulse variation now suggests that the magnitudes of R_1 and R_2 are the same for every specular area which is responsible for the return. The only radius which appears to have an a priori probability of occurring is the moon's radius itself. It is therefore assumed that both radii of curvature are equal to the mean lunar radius, and while this must remain a possible source of error, an alternative choice of radii is discussed in section 7.3.

Replacing R_1 and R_2 by a, the cross section becomes

$$\sigma=\pi a^2|R|^2$$

in agreement with that used in reference [8]. From eqs (15) and (16) we then have

$$\sigma=(d+\lambda^2 d')\pi a^2$$

with $d=4.04\times10^{-4}$ and $d'=2.37\times10^{-5}$ m^{-2}, and the corresponding curve is plotted in figure 5.

7.2. Electromagnetic Constants

The electromagnetic constants of the material of which the specular area is composed can be calculated from the above values of d and d'. If these are inserted into eqs (12) and (13), it is found that

$$\epsilon/\mu=6.5\times10^{-6}\ \text{mhos}^2,$$
$$s/\mu=2.4\times10^{2}\ \text{mhos/henry},$$

or

$$\epsilon/\mu=7.6\times10^{-6}\ \text{mhos}^2,$$
$$s/\mu=2.7\times10^{2}\ \text{mhos/henry}.$$

The two sets appear because of the ambiguity in sign in taking the square root of d. Aesthetically at least, it is more pleasing to select that set for which the ratio ϵ/μ exceeds its value for free space, and this is the second of those given above. We note in passing that only a preliminary version of the first based on seven experimental points was given in references [8] and [24].

From a power measurement alone, ϵ and s cannot be obtained explicitly, but if it is assumed that $\mu=\mu_0$, then

$$\epsilon=9.6\times10^{-12}\ \text{farads/m}(=1.1\epsilon_0)$$
$$s=3.4\times10^{-4}\ \text{mhos/m}.$$

Both of these values are smaller than expected and few (if any) naturally occurring substances on earth, apart from liquids or gases, have a relative permittivity as low as 1.1. Although this is no reason for ruling out the possibility of an appropriate lunar substance, it may be of interest to observe that if μ is increased by a factor α, ϵ, and s are increased by the same amount. Taking α to be, for example, 1.4, we now have

$$\mu=1.4\mu_0,$$
$$\epsilon=1.5\epsilon_0,$$
$$s=4.8\times10^{-4}\ \text{mhos/m},$$

and these are not inconsistent with soils such as magnetite.

7.3. Discussion

Over the range of wavelengths from 0.1 to 2.5 m the variation in the measured cross section is relatively small but is accurately reproduced by the wavelength dependence of the formula for $|R|^2$. In the authors' opinion, this variation is significant and implies a small (but nonzero) conductivity for the material of which the specular area is composed. On the other hand, Purcell [25] has pointed out that a cross section which is wavelength independent can be chosen to fit the data to within ±3 db. A cross section of this type would be represented by a line parallel to the abscissa in figure 5 and would imply either a material of zero conductivity or one whose conductivity varied inversely as the wavelength. For this frequency range at least, the cross section would be approximately 3.5×10^9 m^2.

Purcell has also suggested that in eq (17) the radii of curvature should be about a tenth of the lunar radius, thereby increasing $|R|^2$ by a factor of (approximately) 10^2. This now agrees with the reflection coefficient of certain dry sands on earth, and even with $\mu=\mu_0$ the resulting permittivity is increased to $1.5\epsilon_0$.

Unfortunately, the decreased radius of curvature has several unpleasant consequences. In a short pulse return from the moon, the second major peak is similar to the first in all important respects. To a somewhat lesser extent, this is also true of the third and fourth peaks and, in particular, the power levels do not differ appreciably. Purcell's picture then requires that each peak be produced by a region whose radii of curvature are less than that of the lunoid. Instead of a single bump in the center of the moon, we are now forced to postulate a series of undulations of amplitude $a/10$, and these have to extend over a considerable portion of the lunar surface. It seems unlikely that such a system of (smooth) undulations would have escaped detection in lunar photographs.

If, on the other hand, the radii of curvature are equal to the radius of the moon, the low permittivity may have to be accepted, though the possibility that μ exceeds μ_0 should still be borne in mind. However, the authors are not as yet convinced that a relative permittivity as low as 1.1 is necessarily inconsistent with a solid material in the environment of the moon, and feel that selected powders might conceivably fit the bill.

8. Conclusions

The theory of a quasi-smooth moon was arrived at from a study of many sets of experimental data, and it is felt that the pulse shapes and power returns can be satisfactorily explained on this basis. It is found that the major scattering areas are concentrated near to the center of the moon, and their total number appears to be approximately 25. An indication of their distribution can also be obtained, but no attempt has been made to associate these areas with individual features on the moon. There is clearly scope for further work in this connection.

The area nearest to the earth provides a specular return, and if pulses of short duration are employed, this return shows up as a peak which can be distinguished from the remainder of the pulse. Only a scattering process of this type could produce the high pulse-to-pulse correlation and the lack of wavelength dependence observed in the leading peak. It seems probable that the adjacent scattering areas are also specular in character, but no statement is yet possible about the areas as a whole.

The power level of the leading peak now determines the scattering cross section of the area nearest to the earth, and this in turn specifies the electromagnetic constants for the material of which the area is composed. Using the measured power returns at a variety of different wavelengths the ratios of permittivity to permeability and conductivity to permeability for the surface have been found. The values are smaller than expected, but are not necessarily inconsistent with certain powdered materials in the environment of the moon. In this connection, laboratory experiments under vacuum conditions will be carried out in the near future to determine the constants for several substances whose occurrence on the moon has been postulated.

The analysis reported in this paper and, indeed, the development of the theory itself, were only possible because of the experimental data made available to us. We therefore express our gratitude to the many experimenters who freely communicated their results to us even prior to publication elsewhere. We are also indebted to members of the Radiation Laboratory, among whom should be mentioned R. E. Hiatt, who performed the experiment described in section 3.3, and W. E. Fensler, who provided assistance with this manuscript. The work described herein was carried out for the National Aeronautics and Space Administration under Grant NaG–4–59.

9. References

[1] P. G. Sulzer, G. F. Montgomery, and I. H. Gerks, An uhf moon relay, Proc. I.R.E. **40,** 361 (1952).
[2] J. H. Trexler, Lunar radio echoes, Proc. I.R.E. **46,** 286 (1958).
[3] J. H. DeWitt, Jr., and E. K. Stodola, Detection of radio signals reflected from the moon, Proc. I.R.E. **37,** 229 (1949).
[4] F. J. Kerr and C. A. Shain, Moon echoes and transmission through the ionosphere, Proc. I.R.E. **39,** 230 (1951).
[5] I. C. Browne, J. V. Evans, J. K. Hargreaves, and W. A. S. Murray, Radio echoes from the moon, Proc. Phys. Soc. B **69,** 901 (1956).
[6] A. C. B. Lovell, Report on British work carried out in radio astronomy since 1954, presented at XIIth General Assembly of URSI, Boulder, Colo. (1957).
[7] B. S. Yaplee, R. H. Bruton, K. J. Craig, and N. G. Roman, Radar echoes from the moon at a wavelength of 10 cm., Proc. I.R.E. **46,** 293 (1958).
[8] T. B. A. Senior and K. M. Siegel, Radar reflection characteristics of the moon, Paris Symp. on Radio Astron. (ed. R. N. Bracewell) Stanford Univ. Press (1959).
[9] J. V. Evans, S. Evans, and J. H. Thomson, The rapid fading of moon echoes at 100 Mc/s, Paris Symp. on Radio Astron. (ed. R. N. Bracewell), Stanford Univ. Press (1959).
[10] W. A. S. Murray and J. K. Hargreaves, Lunar radio echoes and the Faraday effect in the ionosphere, Nature **173,** 944 (1954).
[11] B. C. Blevis, personal communications (8 October 1957 and 3 June 1958).
[12] J. Chapman, personal communications (15 to 22 October 1959).
[13] J. S. Hey and V. A. Hughes, Radar observations of the moon at 10 cm wavelength, Paris Symp. on Radio Astron. (ed. R. N. Bracewell), Stanford Univ. Press (1959).
[14] B. S. Yaplee, N. G. Roman, K. J. Craig, and T. F. Scanlan, A lunar radar study at 10 cm wavelength, Paris Symp. on Radio Astron. (ed. R. N. Bracewell) Stanford Univ. Press (1959).
[15] A. B. Youmans, Personal communication (6 June 1958).
[16] J. Aarons, H. Peters, and H. Whitney, Personal communications (29 to 30 October 1959).
[17] V. A. Hughes, Personal communication (28 April 1958).
[18] S. J. Fricker, R. P. Ingalls, W. C. Mason, and M. L. Stone, UHF moon reflections, Presented at URSI–IRE meeting, Washington, D.C. (1958).
[19] R.L. Leadabrand, Personal communication (30 October 1959).
[20] H. D. Webb, Personal communication (30 October 1959).
[21] J. V. Evans, The scattering of radio waves by the moon, Proc. Phys. Soc. B **70,** 1105 (1957).
[22] V. H. Weston, Exact near field and far field solution for the backscattering of a pulse from a perfectly conducting sphere, Univ. of Michigan Radiation Lab. Rept. No. 2778–4–T (1959).
[23] J. A. Stratton, Electromagnetic Theory (McGraw-Hill Book Co., Inc., New York, N.Y., 1941).
[24] W. E. Fensler, T.B.A. Senior, and K. M. Siegel, Exploring the depth of the surface layer of the moon from a radar space observatory, Aero/Space Eng. **18,** 38 (1959).
[25] E. M. Purcell, Personal communication (15 October 1959).

10. Appendix: A Radar Cross Section Theorem

Consider a body on which is incident a field whose direction and polarization are fixed in space. The scattered field is received with two orthogonally polarized antennas which are located at a fixed point in space, but whose polarization is variable. To begin with, it will be assumed that these antennas are in the far field of the body.

At any point in space the scattered field can be written as

$$\underline{E}^s = L'\hat{r} + M'\hat{\theta} + N'\hat{\phi},$$

where $\hat{r}, \hat{\theta}, \hat{\phi}$ are unit vectors for a system of spherical polar coordinates whose origin is in the neighborhood of the body. L', M', N' are functions of the incident field (in direction and polarization), the body (in shape and aspect), and the position of the receiving point. If this point is in the far field of the body, then

$$\lim_{r\to\infty} L' = \mathrm{O}(1/r^2)$$

$$\lim_{r\to\infty} M' = \frac{e^{ikr}}{r} M + \mathrm{O}(1/r^2)$$

$$\lim_{r\to\infty} N' = \frac{e^{ikr}}{r} N + \mathrm{O}(1/r^2)$$

where M and N are independent of r, a time dependence $e^{-i\omega t}$ having been assumed. Hence

$$\underline{\mathrm{E}}^s \sim \frac{e^{ikr}}{r} \underline{\mathrm{F}}$$

with

$$\underline{\mathrm{F}} = M\theta + N\hat{\phi}$$

For a receiving antenna which is polarized in the direction $\hat{x}$, where $\hat{x}$ makes an angle β with some fixed direction, the cross section which is measured is

$$\sigma(\alpha,\beta) = 4\pi|\underline{\mathrm{F}}\cdot\hat{x}|^2$$

and this can be written in the form

$$\sigma(\alpha,\beta) = 4\pi|M\cos\Omega + N\sin\Omega|^2$$

where Ω is the angle between the projection of $\hat{x}$ in the (θ, ϕ) plane and the vector $\hat{\phi}$. Similarly, for the receiving antenna which is orthogonally polarized, the cross section is

$$\sigma\left(\alpha, \beta+\frac{\pi}{2}\right) = 4\pi|M\sin\Omega - N\cos\Omega|^2$$

and hence

$$\sigma(\alpha,\beta) + \sigma\left(\alpha, \beta+\frac{\pi}{2}\right) = 4\pi(|M|^2 + |N|^2). \quad \text{(A.1)}$$

For a fixed transmitted field and for a fixed location of the receiving antennas, the right-hand side of this equation is at most a function of the aspect of the body (which can be changed by rotating the body keeping the antennas fixed). In particular, it is independent of Ω (and therefore β), and in consequence it holds for all receiver polarizations.

If the receiving antennas are not in the far field, only a trivial modification of this proof is required. The field component L' must now be retained, and the equation corresponding to (A.1) is then

$$\sigma(\alpha,\beta) + \sigma\left(\alpha, \beta+\frac{\pi}{2}\right) = \lim_{r\to\infty} 4\pi r^2 (|L'|^2 + |M'|^2 + |N'|^2).$$

To the first order in r this is identical to eq (A.1).

Boulder, Colo. (Paper 64D3–51)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

A Theory of Wavelength Dependence in Ultrahigh Frequency Transhorizon Propagation Based on Meteorological Considerations[1]

Ralph Bolgiano, Jr.[2]

(October 26, 1959)

Recent radio data indicate that the wavelength dependence of ultrahigh frequency transhorizon propagation varies widely in time. This is in contradiction with theoretical explanations previously set forth. Each attempt to account for the underlying effects of ever-present atmospheric motions has, in the past, pointed toward a unique form of the dependence. Extensive discussions have resulted as to the validity and relative merits of the various forms, but at no time has a variable wavelength dependence been proposed.

Since scatter propagation theory has predicted so satisfactorily the broad aspects of the radio signals, it is retained as the basis for further analysis. A new model is developed for the structure of refractive index fluctuations induced by turbulence. Grounded on a theory of homogeneous turbulence in a stably stratified atmosphere, which has been developed concurrently by the author, this new model provides an explanation for the observed distribution of wavelength dependence. It suggests that at times when the *dynamic stability* of the air within the scattering volume is neutral the received power should be nearly independent of radio wavelength. On the other hand, when the atmosphere is dynamically stable the signal strength should be proportional to the square, or higher power, of the wavelength.

These predictions have been tested by comparing the results of a *scaled-frequency* experiment with simultaneous meteorological data gathered along the path. Richardson's number for the 1- to 3-kilometer layer, within which the principal scattering volume lies, has been employed as an index of dynamic stability, though it falls short of ideal in some respects. The 0.8 value of correlation found between Richardson's number and the wavelength dependence is highly suggestive that a relation of the nature predicted does, in fact, exist.

1. Introduction

A central prediction of the Booker-Gordon [1][3] scatter model for extra-diffraction tropospheric propagation was that the scatter loss should be independent of the radiofrequency. This matter of wavelength dependence has received considerable attention over the intervening years [2, 3, 4, 5, 6, 7]. Various other explanations of the underlying electromagnetic phenomenon have been put forth, and several frequency scaling laws have been suggested. Those proposals that have met with greatest favor, both because of their relative simplicity as well as their ability to predict satisfactorily the broad aspects of nearly all the empirical data, are the single-scattering models such as Booker and Gordon [1] employed. Even among the proponents of these models there has been disagreement.

Perhaps this disagreement is most familiar when stated in terms of the spatial correlation of the deviations of refractive index, e.g.,

$$R_n(\underset{\sim}{\xi}) \equiv \overline{\delta n(\underset{\sim}{x},t)\ \delta n(\underset{\sim}{x}+\underset{\sim}{\xi},t)}. \tag{1}$$

It is, however, more readily interpreted when put in terms of the mean square wavenumber spectrum of fluctuations of n within the volume V,

$$\overline{|\delta A_n(\underset{\sim}{k},t)|^2}.$$

$$\delta A_n(\underset{\sim}{k},t) \equiv \int_V \delta n(\underset{\sim}{x},t)\ \exp\{-\imath \underset{\sim}{k}\cdot\underset{\sim}{x}\}\ d\underset{\sim}{x}, \tag{2}$$

$$\frac{\overline{|\delta A_n(\underset{\sim}{k},t)|^2}}{V} = \int_\xi R_n(\underset{\sim}{\xi})\ \exp\{-\imath \underset{\sim}{k}\cdot\underset{\sim}{\xi}\}\ d\underset{\sim}{\xi},$$

the latter integration being over all of $\underset{\sim}{\xi}$-space. This ease of interpretation comes about because the scattering coefficient σ may be expressed directly in terms of this spectrum [3].

$$\sigma = \frac{4\pi^2 \sin^2 \chi}{\lambda^4} \frac{\overline{|\delta A_n(\underset{\sim}{k}_\sigma,t)|^2}}{V}. \tag{3}$$

The wavenumber $\underset{\sim}{k}_\sigma$, at which the spectrum must be evaluated, is determined by the radio wavelength and the path geometry. Its magnitude is given by

$$k_\sigma = \frac{4\pi}{\lambda} \sin\frac{\beta}{2}, \tag{4}$$

in which β is the scatter angle. If the deviations of refractive index are isotropic, at least at scales of

[1] This research was supported by the U.S. Air Force Cambridge Research Center under Contract No. AF 19(604)-3494, with Cornell University.
[2] School of Electrical Engineering, Cornell Univ., Ithaca, N.Y. (invited paper).
[3] Figures in brackets indicate the literature references at the end of this paper.

order corresponding to k_σ, the mean square fluctuation spectrum is a function of wavenumber magnitude only. Then the scattering coefficient may be expressed as

$$\sigma = \frac{8\pi^4 \sin^2 \chi}{\lambda^4} \frac{E_n(k_\sigma)}{k_\sigma^2} \tag{3a}$$

in which

$$E_n(k) = \frac{k^2}{8\pi^3 V} \int_{4\pi} \overline{|\delta A_n(\tilde{k},t)|^2}\, d\Omega(\tilde{k}),$$

the integral extending over all directions of $\tilde{k}$. Thus, the wavelength and angle dependence of the scattering coefficient and, indirectly, the wavelength dependence of the received power stem from the form of this spectrum.

Conversely, to the extent that the single-scattering approximation is valid, the structure of the refractive index variations may be studied experimentally via radio data. This latter approach led Norton [8] to propose his log-normally modified $\rho K_1(\rho)$ correlation function for the n-fluctuations. The $\rho K_1(\rho)$ form had been suggested previously [9] since it leads to a λ to the first power dependence in σ, in agreement with the mean of radio data collected on numerous paths over several years [5]. Villars and Weisskopf [4] suggested a model of turbulent mixing which yields the corresponding spectral form. When their work first appeared, the justification it seemed to afford the empirical results led to its rapid acceptance. More recently, however, an inconsistency has been pointed out in this model of the structure of fluctuation spectra [10]. Wheelon [11] has discussed the matter at some length, but his arguments have been questioned further by Bolgiano [12] and the uncertainty regarding the Villars-Weisskopf mixing theory remains.

Applying the statistical theory of isotropic turbulence, Batchelor [13] has shown that σ should exhibit a $\lambda^{-1/3}$ wavelength dependence. This is a consequence of the $k^{-5/3}$ wavenumber dependence of the mean square refractive index fluctuation spectrum shown in figure 1. This k-dependence is, in turn, the universal form [14] characteristic of the *inertial subrange*, that interval lying between the anisotropic mixing eddies at the large scale end of the spectrum (small k) and the very small size components, subject to severe molecular "dissipation," at the other extreme. In the isotropic theory this dissipation, or viscous, cutoff of the spectrum occurs at a scale η given by

$$\eta = \left(\frac{\nu^3}{\epsilon}\right)^{1/4} \tag{5}$$

where ν is the kinematic viscosity of the air and ϵ is the rate (per unit mass of the fluid) at which turbulent energy is converted into heat by the action of viscosity. On the basis of previous estimates of ϵ that have been made by the meteorologists, η is of the order of millimeters, whereas the scales involved in radio scattering (corresponding to k_σ) are typically of the order of meters. Turbulence theory, therefore, predicting a $\lambda^{-1/3}$ dependence, would appear to be in disagreement with the bulk of the radio data for which the mean form is λ^1.

Partly, no doubt, because of this inability of scatter theory, as based on turbulence-induced dielectric inhomogeneities, to agree with the empirical results, but also because there is strong evidence of refractive index stratification in the lower troposphere, Bauer [15] and Friis, Crawford, and Hogg [7] have proposed coherent partial reflections from layers to explain transhorizon propagation. By judicious choice of the size of the layers (relative to Fresnel zones) and of the vertical thickness of the index gradient modification regions, the latter authors have arrived at a wavelength dependence for the received power of λ^1. It is interesting to note that by other choices (perhaps less sound physically) agreement can be achieved with any wavelength dependence in the range λ^{-2} to λ^4.

In order to account for the signal levels observed, the angular distribution of received power, and the fading character of the signals, these same authors have found it necessary to assume large numbers of layers, randomly located and moving. Viewed in this light, it would seem this reflection model is simply another approach to scattering by turbulent anisotropic irregularities. If so, the appropriate wavelength dependence will stem from the refractive index correlation and spectrum functions. These functions, in turn, should have some sound basis in dynamic meteorology, including atmospheric turbulence, which is seldom, if ever, entirely absent.

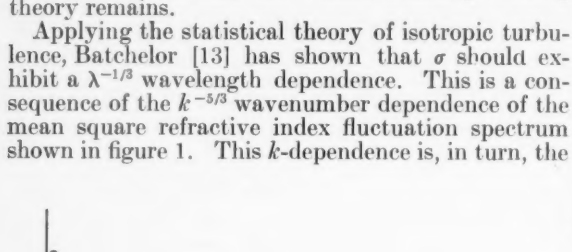


FIGURE 1. *Spectrum of mean square refractive index fluctuations.*
Neutral atmosphere—isotropic turbulence.

2. Distribution of the Wavelength Dependence

The lack of agreement between the predictions of turbulence theory and the propagation data is even more striking when the results of a scaled-frequency experiment [16] are considered. In that experiment, two channels, one at 417 Mc, the other at 2,290 Mc, were operated simultaneously over the same path using scaled antennas. Thus, the volume common

to the transmitting and receiving beams was identical on the two frequencies and any differences in observed transmission loss can be identified directly with the wavelength dependence of σ. From the data of figure 2, it is apparent that no one "universal" spectral form holds, since the wavelength dependence, as indicated by the corrected difference in the observed values of hourly median basic transmission loss on the two channels (ΔL), varies widely.[4] True, the median of the distribution corresponds roughly to λ^1. The variation, however, ranges from $\lambda^{-3/2}$ to $\lambda^{5/2}$, i.e., from $k^{-1/2}$ to $k^{-9/2}$ for the fluctuation spectrum. This then must disqualify any theoretical explanation in terms of a model which predicts a unique wavenumber dependence. Yet turbulence theory, as it has been applied to the problem, has led to such a result.

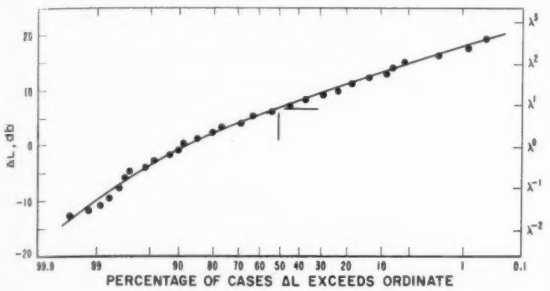


FIGURE 2. *Distribution of wavelength dependence.*

Corrected differences in measured hourly median basic transmission loss, ΔL for scaled-frequency experiment: 417.05 Mc versus 2,290 Mc, Round Hill—Crawford's Hill, February 11, 1957 to July 11, 1957. 241 cases. (After Chisholm, Roche, and Jones [16].)

This seeming dilemma could be overcome if only the viscous cutoff scale η lay in the meter range much of the time, rather than in the millimeter range. Consideration of figure 1 will aid in visualizing this point. If k_d were to fall at or below k_σ, the wavenumber dependence of the spectrum in the vicinity of k_σ would be "stronger" than the $k^{-5/3}$ form of the inertial subrange. Wide variation of η in time could then account for the distribution of wavelength dependence of the received power which has been observed.

These large values of η, however, could occur only in conjunction with exceedingly small dissipation rates; and very small ϵ, in turn, requires either a correspondingly low turbulent intensity or a more rapid decrease of kinetic energy density with increasing wavenumber than the universal equilibrium theory [18] indicates. This latter solution implies the progressive removal of kinetic energy from the spectrum by some mechanism other than viscosity.

Very low turbulent intensity, which may result when a given field of turbulence has decayed appreciably, is certainly a prerequisite to the existence of well-formed layers of the type essential to Bauer's [15] coherent reflection model. When such circumstances prevail, an explanation based instead on turbulence theory (modified as above) has great merit since it does not require a complete absence of random air motions. It states only that such motions must be very "weak." Even this limited requirement, though it may be satisfied occasionally (in a strong inversion for example), is not the typical situation at the 5,000-ft level. This conclusion is borne out by the meteorological evidence as to the usual intensity of turbulent velocities at this altitude [19]. It is several orders of magnitude higher than that of the low level turbulence needed if this explanation is to be applicable.

3. Turbulence in Stably Stratified Media

It should be noted, on the other hand, that the meteorological estimates of ϵ invariably have been made on the basis of some large-scale turbulence effect, such as the modification of the wind profile by eddy viscosity, together with the assumption that the dissipation rate is determined solely by this same large-scale turbulence. This is equivalent to assuming that the rate at which kinetic energy is transferred "up the spectrum" inertially, from larger to smaller eddies, is constant, independent of eddy size, until actual dissipation begins to take effect in the vicinity of the viscous cutoff. That is to say, no mechanism other than viscosity operates to remove kinetic energy from the turbulent hierarchy.

This approach does not adequately account for the effects of buoyancy forces which come into play in a stable atmosphere as a result of the less than adiabatic temperature lapse rate. A particle of air on moving vertically is no longer in static equilibrium with its surroundings. Continuing to travel up or down it works against the force of buoyancy and, as a consequence, has at least part of its (turbulent) kinetic energy converted to potential energy. Vertical flux of heat effected by eddy diffusion is a closely associated process and a familiar meteorological phenomenon.

An important result, therefore, of a stable density gradient is to provide the means by which kinetic energy may be extracted from the velocity field at various intermediate scales, thus voiding the assumption on which the estimates of ϵ depend. It is conceivable that for stable conditions these estimates have been high, perhaps by many orders of magnitude.

In a neutral atmosphere,[5] as in a uniform fluid, there being no preferred direction, it is not suprising that the smaller scale motions have a strong tendency toward isotropy. On the contrary, a stable temperature (and density) gradient must single out the vertical direction as unique. Consequently, an anisotropic, axisymmetric structure will be favored.

Under such circumstances the energy spectrum tensor takes the general form

[4] The is the question as to the statistical significance of this variation. This matter as been considered, and it has been found that the standard deviation of the empirical data is an order of magnitude larger than would be expected on the basis of statistical fluctuations alone [17].

[5] An atmosphere in which the mean potential temperature $\bar{\vartheta}$ is independent of height.

$$\tfrac{1}{2}\Phi_{ii}(\underset{\sim}{k}) = -k^2 A(k) + \tfrac{1}{2} B(k^2, k_z^2)\left(1 + \frac{k_z^2}{k^2}\right), \quad (6)$$

in which A and B are arbitrary scalar functions, and the total energy density may be written

$$E(k) = -4\pi k^4 A(k) + \frac{k^2}{2}\int_{4\pi} B(k^2, k_z^2)\left(1 + \frac{k_z^2}{k^2}\right) d\Omega(\underset{\sim}{k}). \quad (7)$$

Moreover, the spectral covariance between vertical velocity and density fluctuations $\Phi_{z\tilde{\omega}}(\underset{\sim}{k})$, which is identically zero in an isotropic field, is of the form

$$\Phi_{z\tilde{\omega}}(\underset{\sim}{k}) = F(k^2, k_z^2)\left(1 - \frac{k_z^2}{k^2}\right) \quad (8)$$

Since this velocity-density covariance plays an essential role in any interaction of the various eddies with the density gradient it seems likely that such interaction does occur, thereby tending to preserve the anisotropic structure. On the hypothesis that a state of equilibrium of this nature exists among those motions not too large in scale, it is possible to construct a theory of homogeneous turbulence in stably stratified media [20] much like the now familiar universal equilibrium theory for uniform fluids. Use is made of the methods of dimensional analysis, the determining parameters being

ϵ = the kinetic energy dissipation rate,

ν = the kinematic viscosity,

$\mathscr{D}$ = the molecular diffusivity,

$g/\hat{\rho}_a$ = the ratio of acceleration due to gravity to the reference potential density,

ϵ_n = the rate of "dissipation" of mean square refractive index fluctuations,

$\epsilon_{\tilde{\omega}}$ = the rate of "dissipation" of mean square potential density fluctuations.

A prominent feature of this new model is an equilibrium range which, in general,[6] divides into three parts: The *buoyancy subrange*, the inertial subrange, and the dissipation subrange. The buoyancy subrange is that portion at the large scale end in which buoyancy forces play a dominant role and the structure is anisotropic. It is within this interval that kinetic energy is extracted by conversion to potential energy. The rate at which kinetic energy is transferred inertially "up the spectrum" is here many times the dissipation rate ϵ. As a consequence, the structure depends upon only $\epsilon_{\tilde{\omega}}$, $(g/\hat{\rho}_a)$, and ϵ_n. The energy and refractive index fluctuation spectra are given by

$$E(k) \sim \epsilon_{\tilde{\omega}}^{2/5} (g/\hat{\rho}_a)^{4/5} k^{-11/5}, \quad (9)$$

$$E_n(k) \sim \epsilon_n \epsilon_{\tilde{\omega}}^{-1/5} (g/\hat{\rho}_a)^{-2/5} k^{-7/5}. \quad (10)$$

The inertial subrange lies in the middle of the equilibrium range, is dominated solely by the inertial transfer processes, and is isotropic. At these scales the anisotropy of the buoyancy subrange has been erased largely through the nondirectional effects of the pressure fluctuations. The spectral forms are those of the constant density theory, viz,

$$E(k) \sim \epsilon^{2/3} k^{-5/3}, \quad (11)$$

$$E_n(k) \sim \epsilon_n \epsilon^{-1/3} k^{-5/3}. \quad (12)$$

The dissipation subrange contains the smallest scales (highest wavenumber components). It is in this region that the great bulk of viscous (molecular) dissipation occurs.

Figure 3 indicates the nature of E_n in a typical

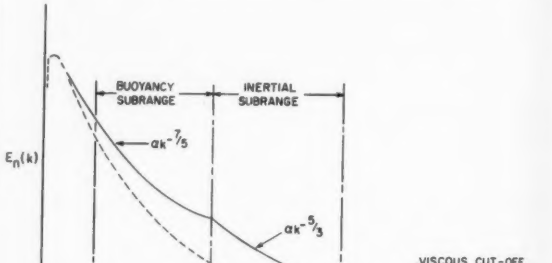


FIGURE 3. *Spectrum of mean square refractive index fluctuations.*

Stably stratified atmosphere—anisotropic turbulence.

case. The dashed curve shows the spectrum for a neutral atmosphere. The effect on the location of the viscous cutoff at k_d should be noted in particular.

The buoyancy and inertial subranges are separated by a wavenumber k_B given by

$$k_B \equiv \frac{\epsilon_{\tilde{\omega}}^{3/4} (g/\hat{\rho}_a)^{3/2}}{\epsilon^{5/4}}. \quad (13)$$

$\epsilon_{\tilde{\omega}}$, being the rate at which mean square potential density fluctuations are dissipated by molecular diffusion, is also the rate at which the same quantity is generated through mixing of the mean potential density gradient by the large scale motions. Its value is therefore determined by the static stability of the atmosphere and the intensity of the primary mixing motions. On the other hand, ϵ depends not only upon the large scale velocity components but also upon the total rate at which kinetic energy is extracted from the turbulence in the buoyancy subrange. Thus, in a neutral or near neutral atmosphere k_B is very small (large size eddy). There is then little or no buoyancy subrange; ϵ is determined by the large scale motions alone, and the universal equilibrium theory results eq (11) and (12) hold. In a highly stable situation, however, especially if there is relatively little mechanical excita-

[6] At the very high Reynolds numbers that characterize the atmosphere.

tion,[7] k_B is very large and the inertial subrange may be entirely suppressed. Under these circumstances, ϵ being exceedingly small (even though the "turbulent intensity" observed at the large scales is moderate), the viscous cutoff may occur at eddies of the order of a meter in size.

The exact form that the spectra take in the dissipation range, beyond the viscous cutoff, is an unsettled matter. Batchelor [21] and Batchelor, Howells, and Townsend [22] have shed some light upon the limiting cases, $\mathscr{D} << \nu$ and $\mathscr{D} >> \nu$. The case of primary interest for the troposphere ($\mathscr{D} \sim \nu$) has yet to be treated satisfactorily.

Before leaving this subject it is interesting to note that there is some evidence, quite independent of the radio propagation phenomenon, which favors this stably stratified model. One of the predictions it yields is that the dispersion of a cloud of particles (a smoke puff, for example) is characterized by a phase of its development in which the diffusion rate increases as the fourth power of time. This compares with a second power form on the basis of the universal equilibrium theory and a constant rate for molecular diffusion. Analysis of a series of smoke puffs in the upper troposphere and lower stratosphere has shown that under conditions of near neutral stability and/or high wind and shear the diffusion rate is closely proportional to t^2 [23], while under conditions of high stability with light wind and weak shear it is more nearly proportional to t^4 (Bolgiano, unpublished). The available data are too few to be conclusive. Nonetheless, such evidence is certainly not contradictory.

4. Relation of Wavelength Dependence to Dynamic Stability

In an attempt to understand more fully the distribution of wavelength dependence, the propagation data of Chisholm, Roche, and Jones [16] and, when available, simultaneous meteorological data for the vicinity of the path have been scrutinized and compared [24]. For example, the winter (February and March) and summer (June and July) data have been plotted separately in figure 4. Note that the standard deviation is larger and the median value higher in summer than in winter. Though the variation from season to season appears distinct, there is no immediately obvious explanation. A similar conclusion may be reached regarding the diurnal cycle. The data have been replotted on a diurnal basis in figure 5. There is no pronounced diurnal variation in winter and even in summer only the lower limit shows a fairly well developed cycle.

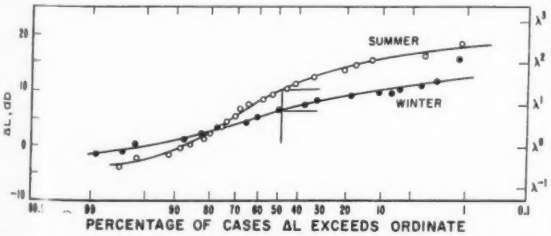


FIGURE 4. *Seasonal variation in distribution of wavelength dependence.*

Corrected differences in measured hourly median basic transmission loss, ΔL, for scaled-frequency experiment: 417.05 Mc versus 2,290 Mc, Round Hill—Crawford's Hill. Winter: February and March 1957, 76 cases. Summer: June and July 1957, 80 cases.

[7] Light winds and little or no shear.

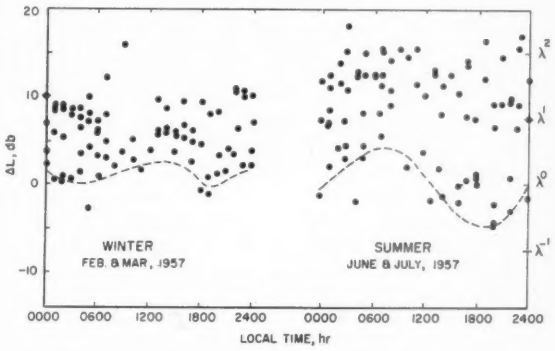


FIGURE 5. *Diurnal variation of wavelength dependence.*

Corrected differences in measured hourly median basic transmission loss, ΔL for scaled-frequency experiment: 417.05 Mc versus 2,290 Mc, Round Hill—Crawford's Hill.

Another prominent feature of the data is the coincidence of abrupt negative changes in the differential transmission loss ΔL, with the occurrence of precipitation along the path. This is particularly noticeable when the radio data are "continuous" for several days and local or frontal showers cross the path. Finally, correlation between the wavelength dependence and the static stability of the atmosphere ($d\bar{\theta}/dz$) has been found significant at the 0.05 level.

This gradient of mean potential temperature $\bar{\theta}$ is indicative of the strength of the buoyancy forces, which operate to oppose vertical motions when the gradient is positive, to accelerate them when it is negative. Its value depends upon the actual temperature gradient and upon the adiabatic temperature lapse rate of the atmosphere. This latter quantity, in turn, changes abruptly as the air becomes saturated, the change being such that less stable conditions result. The precipitation effect noted above could be accounted for on this basis. At least it is now possible to explain the above coincidence in terms of the observed correlation with static stability.

This correlation, as well as the seasonal and diurnal variation depicted in figures 4 and 5, seem to be in qualitative agreement with the notions developed in sections *B* and *C*. In summer when mean winds in the Long Island area [8] are "light-to-variable" (though there are occasional intense thunderstorms) the median value of ΔL is high and the spread is

[8] The Round Hill-Crawford's Hill path, on which the radio data were gathered, is such that the major scattering volume is over Long Island and Long Island Sound.

wide. In winter, on the contrary, when the mean winds are more uniformly intense the median power ratio is lower and the spread is narrower. Again in winter when there is relatively little consistent diurnal variation of weather (mean winds, temperature gradients, etc.) there is no pronounced daily cycle. In summer, however, when there is sometimes marked variation of wind, convection, and turbulence, from calm or light in the early morning hours to moderate or strong in the afternoon, one finds a fairly evident diurnal cycle.

Though such qualitative agreement with the proposed model is encouraging, it is desirable to test the predictions which can be made on the basis of that model in some more quantitative manner. Explicitly these predictions are: (1) That at times when mechanical excitation in the atmosphere (large scale horizontal motions, shear, etc.) dominates stability suppression, the wavelength dependence should favor a $\lambda^{-1/3}$ to $\lambda^{-3/5}$ form; and (2) at times of high static stability, however, and when winds are light, the wavelength dependence should tend toward λ^1 and λ^2 and higher powers.

The desired test might be achieved by investigating the correlation, if any, between an appropriate index of dynamic stability and the differential transmission loss, ΔL. The choice, or formation, of a suitable index which can be evaluated readily from the generally available meteorological data presents some difficulty. Richardson's number Ri[9] meets the specifications only approximately, at best. It fails to account properly for the turbulence-driving capabilities of very large scale horizontal motions.

Nonetheless, Ri has been evaluated from radiosonde and rawinsonde data gathered in the vicinity of the path for the layer 1 to 3 km which includes the primary scattering volume. Mean winds at each level have been calculated and the mechanical excitation taken as the square of the total variation of mean wind over the chosen height interval. The resulting index values have been compared with ΔL for a number of "consistent" cases. These are instances for which the value of ΔL remained reasonably constant (within several decibels) for a minimum of 8 hr, although the transmission loss on the individual channels varied sometimes by 10 to 15 db. The data are presented in table 1.

The scatter plot in figure 6 indicates that the correlation is very pronounced. The value of sample correlation computed is 0.80 which, for this 20-point sample, establishes the lower (one-sided) 95-percent confidence limit, for the true correlation, at 0.6. That is to say, if the actual correlation is less than 0.6, the sample correlation should be less than the observed value with a probability of 0.95, or greater.

[9] Ri is the ratio of the rate at which kinetic energy is extracted from the turbulence, by the effect of buoyancy forces, to the rate at which kinetic energy is transferred to the turbulence from the mean flow through interaction with the vertical wind shear.

TABLE 1. *Data and computed results for "consistent" cases*

Round Hill—Crawford's Hill Radio Link, February-July 1957. Meteorological data from, "daily series, synoptic weather maps, part II: northern hemisphere data tabulations"

Case	ΔL	Precipitation	$\overline{(dT/dz)}_{1-3}$	$\overline{T}_{1-3}$	dT/dz|adiab.	$\|\Delta U\|_{1-3}$	$(Ri)_{1-3}$
Date and time	*db*		°*C*/*km*	°*K*	°*C*/*km*	*m*/*s*	
I (2/28/1500–3/01/0500)	1	▲∴	−2.9	273	−6.0	18.0	1.4
II (3/05/2200–3/06/0700)	9		−3.0	268	−9.9	9.0	13.0
III (3/18/1800–3/19/0600)	3	⊤	−3.8	271	−9.5	21.0	2.1
IV (5/06/2000–5/07/0400)	−11		−9.5	272	−9.9	9.5	0.7
V (5/07/0600–5/07/1800)	2		−6.8	278	−9.9	13.5	2.4
VI (5/13/1200–5/13/2200)	6		−6.0	283	−9.9	15.0	2.5
VII (5/14/0000–5/14/1400)	9		−5.3	282	−9.9	17.5	2.1
VIII (5/21/1800–5/22/0400)	6		−2.4	277	−9.9	11.0	9.0
IX (5/22/0800–5/22/1800)	6		−1.1	282	−9.9	13.0	7.4
X (5/23/0000–5/23/0800)	5	⊤	−3.6	284	−9.5	10.5	7.9
XI (6/11/0100–6/11/1200)	14		−4.7	283	−9.9	9.0	9.0
XII (6/11/1200–6/11/2300)	15		−5.2	285	−9.9	5.5	21.2
XIII (6/24/1300–6/24/2200)	−1	⊼∞	−6.3	286	−7.3	8.0	2.2
XIV (6/26/0700–6/26/1500)	14		−5.2	285	−9.9	7.5	12.1
XV (6/26/2000–6/27/0800)	11		−5.6	285	−9.9	7.5	10.7
XVI (7/02/1300–7/03/0000)	−2		−7.5	280	−9.9	10.5	3.1
XVII (2/11/2200–2/12/0600)	8		−2.6	263	−9.9	12.0	7.8
XVIII (5/15/0600–5/15/1800)	7	⊤	−6.3	283	−9.5	10.5	3.9
XIX (7/08/1600–7/09/0000)	15		−7.5	284	−9.9	8.5	4.6
XX (7/10/1600–7/11/0600)	10		−5.5	278	−9.9	10.5	5.7

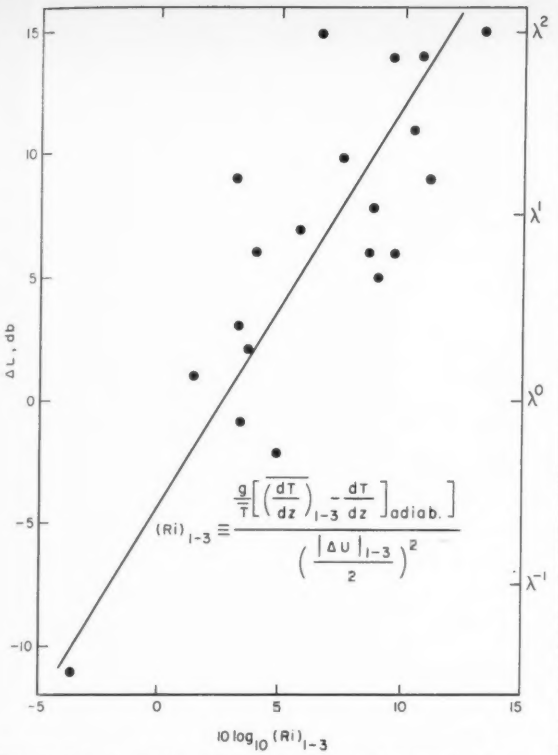


FIGURE 6. *Variation of wavelength dependence with dynamic stability of the atmosphere in the scattering volume.*

Sample correlation $r=0.80$.

5. Conclusions

A rather close relationship appears to exist between the wavelength dependence of transhorizon tropospheric propagation and the dynamic stability of the atmosphere in the region of space common to the transmitting and receiving antenna beams. Considered in terms of scatter theory this implies an association between the functional form of the spectrum of mean square refractive index fluctuations and dynamic stability. This in turn, when viewed in the light of statistical turbulence theory, seems to indicate a strong connection between the viscous cutoff scale and this meteorological index. Finally then, there appears to be in the radio data confirmation of the proposed model of turbulence in stably stratified media including the notion of progressive extraction of kinetic energy from the turbulence and the resultant vast reduction in the dissipation rate ϵ.

The author expresses his gratitude to W. E. Gordon for many helpful discussions and suggestions during the conduct of this investigation; and to the Lincoln Laboratory, Massachusetts Institute of Technology, for so kindly inviting the author to visit its Tropospheric Propagation Group in order that he might analyze the results of certain of its radio experiments.

6. References

[1] H. G. Booker and W. E. Gordon, Theory of radio scattering in the troposphere, Proc. Radio Wave Propogation Symp., Navy Electronics Lab., San Diego, Calif. (1949); also Proc. IRE **38,** 401 (1950).

[2] E. C. S. Megaw, Scattering of electromagnetic waves by atmospheric turbulence, Nature **166,** 1100 (1950).

[3] F. Villars and V. F. Weisskopf, The scattering of electromagnetic waves by turbulent atmospheric fluctuations, Phys. Rev. **94,** 232 (1954).

[4] F. Villars and V. F. Weisskopf, Scattering of radio waves by turbulent fluctuations of the atmosphere, Proc. IRE **43,** 1232 (1955).

[5] K. A. Norton, P. L. Rice, and L. E. Vogler, Use of angular distance in estimating transmission loss, Proc. IRE **43,** 1488 (1955).

[6] R. A. Silverman, Turbulent mixing theory applied to radio scattering, J. Appl. Phys. **27,** 699 (1956).

[7] H. T. Friis, A. B. Crawford, and D. C. Hogg, A reflection theory for propagation beyond the horizon, Can. Conv. IRE, Toronto, Canada (1956); Bell System Tech. J. **36,** 627 (1957); also A. B. Crawford, D. C. Hogg, and W. H. Kummer, Bell System Tech. J. **38,** 1067 (1959), for a further discussion of the variability of wavelength dependence as influenced by layer size relative to a subtended Fresnel zone. (Ed.)

[8] K. A. Norton, Point-to-point radio relaying via the scatter mode of tropospheric propagation, IRE Trans. **CS–4,** 39 (1956); unpublished report (1955), in particular the Note added July 28, 1956, introducing the log-normal modification.

[9] R. B. Muchmore, private communication to K. A. Norton (1955), reference [8].

[10] R. Bolgiano, Spectrum of turbulent mixing, Phys. Rev. **108,** 1348 (1957).

[11] A. D. Wheelon, On the spectrum of a passive scalar mixed by turbulence, J. Geophys. Research **63,** 849 (1958).

[12] R. Bolgiano, Jr., On the role of convective transfer in turbulent mixing, J. Geophys. Research **63,** 851 (1958).

[13] G. K. Batchelor, The scattering of radio waves in the atmosphere by turbulent fluctuations in refractive index (Cornell Univ., School of Elec. Eng. Research Rept. **EE 262,** Ithaca, N.Y., 1955).

[14] R. Bolgiano, The role of turbulent mixing in scatter propagation, IRE PGAP Trans. **AP–6,** 161 (1958).

[15] J. R. Bauer, The suggested role of stratified elevated layers in transhorizon short-wave radio propagation (Lincoln Lab., Tech. Rept. No. 124, Bedford, Mass., 1956).

[16] J. H. Chisholm, J. F. Roche, and W. J. Jones, Experimental investigations of angular scattering and multipath delays, Proc. of Combined Tech. Session, URSI–IRE, Washington, D.C. (1957).

[17] R. Bolgiano, Jr., Wavelength dependence in transhorizon, propagation, Proc. IRE **47,** 331 (1959).

[18] G. K. Batchelor, The theory of homogeneous turbulence (Cambridge Univ. Press, London, England, 1953).

[19] H. Charnock, Note on eddy diffusion in the atmosphere between one and two kilometers, Quart. J. Roy. Meteorol. Soc. **77,** 654 (1951).

[20] R. Bolgiano, Jr., Turbulent spectra in a stably stratified atmosphere, J. Geophys. Research **64,** 2226 (1959).

[21] G. K. Batchelor, Small-scale variation of convected quantities like temperature in turbulent fluid, pt. 1: General discussion and the case of small conductivity, J. Fluid Mech. **5,** 113 (1959).

[22] G. K. Batchelor, I. D. Howells, and A. A. Townsend, Small-scale variation of convected quantities like temperature in turbulent fluid, pt. 2: The case of large conductivity, J. Fluid Mech. **5,** 134 (1959).

[23] F. Gifford, Jr., Relative atmospheric diffusion of smoke puffs, J. Meteorol. **14,** 410 (1957).

[24] R. Bolgiano, Jr., A meteorological interpretation of wavelength dependence in transhorizon propagation (Cornell University, School of Elec. Eng. Research Rept. **EE 385,** Ithaca, N.Y., 1958).

BOULDER, COLO., (Paper 64D3–52)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

A Preliminary Study of Radiometeorological Effects on Beyond-Horizon Propagation

F. Ikegami [1]

(October 22, 1959)

A study was made of American and Japanese radiometeorological data in order to suggest the dominant factors in propagation beyond the horizon. The diurnal variability of radio field strengths seems to be sensitive to the crossover height of rays tangent to the radio horizon and disappears for crossover heights greater than about 500 meters. High hourly median field strengths were observed in Japan corresponding to the existence of a marked refractive index discontinuity layer in a common volume of two antenna beams. The results of these experiments suggest that laminar structures of the atmosphere play an important role in beyond-horizon radio propagation.

1. Introduction

Detailed investigations have established the correlations between monthly median values of basic transmission loss and radiometeorological factors. These correlations have been successfully applied to design practical communication links utilizing the surface refractivity or the lapse rate of refractivity with height in the first kilometer above the earth's surface [1 2 3 4].[2] The next stage of research requires the estimation of the variation of transmission loss within a month in order to determine the nature of the fading in more detail.

Several theories have been proposed for the explanation of low-transmission loss beyond the horizon unexpected by classic diffraction theory; for example, scattering [5], partial reflection in the tapered atmosphere [6], diffraction [7], reflection at multiple layers [8], reflection at an elevated layer, normal refraction, and superrefraction. Usually the first four stand for lower signal strength, while the last three, for the higher. None of them may be consistent for actually observed propagation phenomena because the observed field is apparently dependent upon various meteorological structures of the atmosphere with different modes of propagation.

However, the solution of some engineering problems requires a knowledge of the main factors which effectively predominate in actual propagation phenomena rather than the determination of the propagation mechanism to which the lowest field intensity beyond the horizon (so-called "scattered" level) is attributed.

This paper describes a preliminary study of the meteorological effects on beyond-horizon propagation, based upon the experimental results obtained by NBS and the Electrical Communication Lab. of Nippon Telegraph and Telephone Corp., Japan (ECL), and presents two aspects that should be taken into consideration for further studies towards practical application. The first aspect concerns the well-known diurnal variation of transmission loss over land; the second concerns low elevated layers that appear from time to time in the atmosphere.

Some considerations of the experimental relations between radio and meteorological data are made and suggestions offered for further studies in this field.

2. Diurnal Variation of Hourly Median Transmission Loss Beyond the Horizon

2.1. Description of Experimental Results (NBS)

The results of the extensive propagation experiments conducted by NBS show that the transmission loss to points near and just beyond the horizon is subjected to a systematic diurnal change, and that the diurnal variation disappears far beyond the horizon. The dependence of diurnal variation on angular distance has been reported in detail [1, 9]. We will examine the diurnal variation from another standpoint. In order to compare diurnal variations for various propagation paths, the period of August 1952 was chosen mainly because measurements of transmission loss were carried out simultaneously or in similar meteorological conditions for the waves radiated at Cheyenne Mt., Camp Carson, and Pikes Peak. The lengths of the periods of data used in this analysis were not always long enough to completely express all diurnal characteristics, since some are even of insufficient length for analysis. All data were used here, however, in order to complement the shortage of varieties in propagation paths.

In figure 1 is shown the mean diurnal variation of hourly median transmission loss for various propagation paths in terms of increasing height of cross-over point of the two horizons above the ground. The height of the cross-over point is used in this paper instead of angular distance in order to discuss its connection with meteorological effects on diurnal

[1] Electrical Communication Lab., Nippon Telegraph and Telephone Corp., Tokyo, Japan, on leave as a United Nations Technical Assistance Fellow.
[2] Figures in brackets indicate the literature references at the end of this paper.

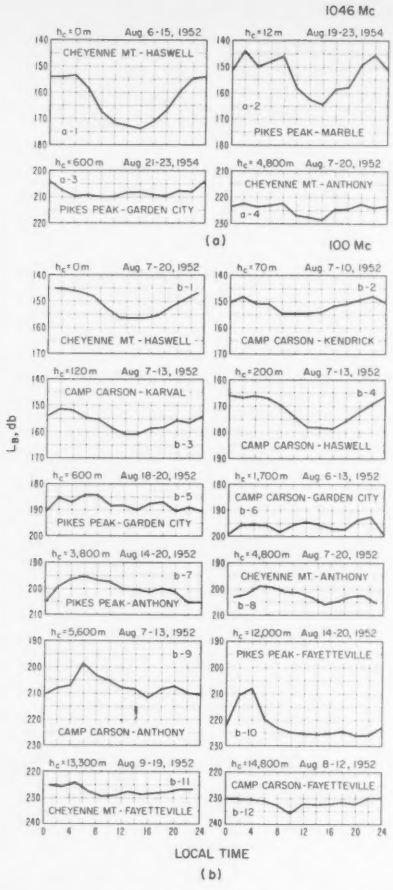


FIGURE 1(a) and 1(b). *Mean diurnal variations of hourly median transmission loss.*

variation of transmission loss. The height of the crossover point, of course, corresponds to angular distance with a one-to-one correspondence, if a spherical smooth earth is assumed. It is easily seen from figure 1 that systematic diurnal characteristics rapidly disappear with increasing height of the crossover point, and that amplitudes of systematic diurnal variation depend on radiofrequency and have a maximum just beyond the horizon, as pointed out in reference [1]. We will see that the unusual increases of field strength seen on paths between Pikes Peak and Fayetteville and between Camp Carson and Anthony do not stand for systematic diurnal characteristics but rather are due to unusual signal enhancement that took place once on each path during the test period.

We now wish to examine the dependence of systematic diurnal characteristics of the transmission loss upon the height of the crossover point. The amplitude of the diurnal variation itself has an important meaning in practical engineering and should be analyzed in a further stage of research based upon a large quantity of data obtained on various propagation paths and in various seasons, and/or upon the propagation mechanism, if clarified, in relation with pertinent radiometeorological factors.

Several methods might be available to express the importance of systematic diurnal characteristics in the total variation of transmission loss. A rough but simple method is applied here as a trial. We assume that the total range ΔL of the hourly median transmission losses consists of two components, ΔL_d and ΔL_n, the former being subjected to systematic diurnal variation and the latter to nonsystematic variation. Then

$$V_d=\frac{\Delta L_d}{\Delta L}=\frac{\Delta L_d}{\Delta L_d+\Delta L_n} \tag{1}$$

represents the importance of diurnal variation in the total variation, which may be temporarily referred to as "diurnal variability" hereafter. Noting that maximum and minimum signal strengths are expected at 02 and 14 hr of the day, respectively, in an ordinary diurnal pattern, and considering that radiosonde observation data at Denver are available at these times, the difference of transmission loss in decibels at 02 and 14 hr divided by the maximum variation of transmission loss observed during the test period is used as an expression of diurnal variability.

That is:

$$V_d=\frac{(\overline{L}_B)_{14}-(\overline{L}_B)_{02}}{(L_B)_{\max}-(L_B)_{\min}} \tag{2}$$

where

V_d=diurnal variability (≤ 1),
$(\overline{L}_B)_{14}$=mean of hourly median transmission losses at 14^h,
$(\overline{L}_B)_{02}$=mean of hourly median transmission losses at 02^h,
$(L_B)_{\max}$=maximum hourly median basic transmission loss during the test period,
$(L_B)_{\min}$=minimum hourly median basic transmission loss during the test period.

This value is not an expression of the absolute importance of diurnal variation but is merely a relative one. Comparison of diurnal variability, therefore, should be made for the same or similar meteorological conditions. A large value of V_d means that a large portion of the total variation is caused by factors subjected to a systematic diurnal characteristic, and vice versa.

The variation of diurnal variability with crossover height is shown in figure 2 for the propagation paths mentioned above, as well as for that between Cedar Rapids and Quincy [9]. Except for one point of the Cheyenne Mt.—Anthony path at 1,046 Mc, plotted points show a clear tendency for V_d to be larger the lower the crossover height, h_c. Further, V_d decreases rapidly as h_c increases. The value of V_d seems to

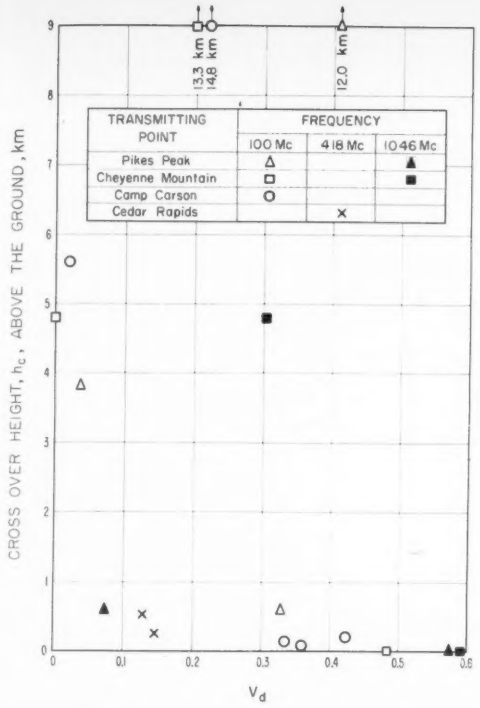


FIGURE 2. *Variation of diurnal variability with crossover height.*

increase at extreme heights of crossover point. These extreme points are all for Fayetteville, and correspond to the case described above when an unusual increase in field strength occurred once during a short test period. This result indicates that a systematic diurnal variation appears when the crossover height is located below about 500 m above the ground and diminishes almost completely above about 1,000 m.

It is interesting to note that a similar tendency of V_d is seen regardless of the height of the transmitters; for example, the transmitter height was very much larger on the Pikes Peak path than on the Camp Carson and Cedar Rapids paths. This shows the significance of crossover height, rather than either that of distance or of antenna height, in the expression of diurnal variation of transmission loss.

2.2. Considerations of Diurnal Variation

Out of several modes of propagation beyond the horizon, a combination of diffracted and scattered field seems a most probable explanation of diurnal variation. However, this fails to explain the large amplitude of diurnal variation by using the ordinary linear N-profile concept [1].

Another interesting phenomenon should be pointed out. Field strengths recorded within an hour on both the Camp Carson—Haswell and Cedar Rapids– Quincy paths are frequently subjected to Rayleigh distributed fading. This even occurs during the night when median field intensities are higher by 10 or more decibels than in the daytime. We find a difficulty in trying to explain this phenomenon by means of an ordinary combination of diffraction and scattering modes with only refraction in the lowest part of the atmosphere superimposed. The exact answer for this question is very difficult to obtain so far, as measurements are not yet available for the fine structure of the lower atmosphere.

At this point we are reminded of the dependence of diurnal variability on crossover height. It is well known that the lower part of the atmosphere is deformed by solar radiation and nocturnal cooling with a distinguished diurnal cycle, and very much has been discussed about diurnal variation in the structure of the lower atmosphere. It is interesting to note that diurnal variation of the atmosphere is limited to the lowest part, ranging from the earth's surface up to 500 m but perhaps not higher than 1,000 m above the ground. For example, according to the extensive observations at Larkhill and Downham Market [10], mean heights of temperature inversion layers are 200 to 300 m depending upon season and time of the day. Similar results were obtained in radiometeorological observations performed by ECL [11] utilizing a 300 m high tower. This latter experiment showed that temperature inversion layers extending up to 200- or 300-m high commonly displayed a diurnal cycle, and the gradient of refractive index versus height also showed a remarkable diurnal variation. Since the exact mechanism of radio propagation has not been clarified yet, no specific meteorological factor could be discussed here with particular significance. However, the above results may be compared with figure 3, which shows the diurnal variation of N-gradient at various heights above the ground observed in Denver during the radio test period in August 1952. Differences of N-gradient at 02^h and 14^h, plotted in figure 4 against height, indicate that diurnal variation in the refractive index structure of the atmosphere rapidly decreases with height and disappears above about 500 m.

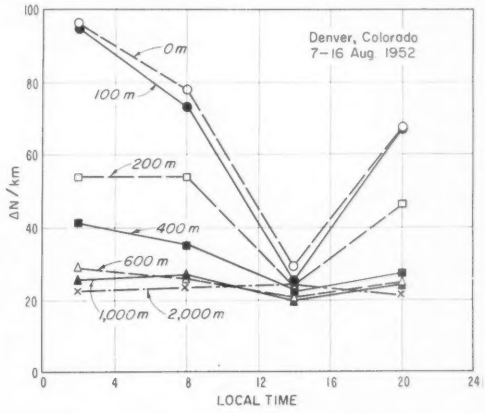


FIGURE 3. *Diurnal variations of the gradient of refractivity at various heights above the ground.*

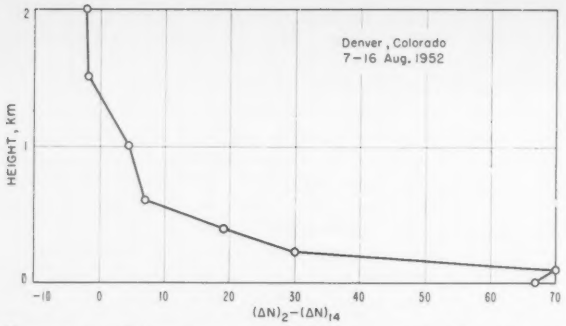


FIGURE 4. *Diurnal range of* ΔN/km *versus height above the ground.*

Though a comparison of figures 2 and 4 is not based on any strict physical reason, the similarity shown in these figures might suggest the large contribution of the atmosphere near the crossover point of the path to propagation beyond the horizon. In fact, the difficulty may be reduced by assuming that the field strength beyond the horizon is largely dependent upon the refractive index structure of the atmosphere, for instance, laminar structure in the lower part of the common volume of two antenna beams, considering that the lowest part of the atmosphere may contribute to emphasize it by lowering the crossover height.

3. Effects of an Elevated Layer in Beyond-Horizon Propagation

3.1. Experimental Results (ECL)

The Radio Wave Propagation Section of Electrical Communication Lab., N.T. & T., Japan, conducted a series of beyond-horizon oversea propagation tests from the fall of 1957 until the summer of 1958 [12]. The map and terrain profile of propagation paths of Miyazaki—Muroto and Miyazaki—Wakayama are shown in figures 5 and 6. In the test of May 1958,

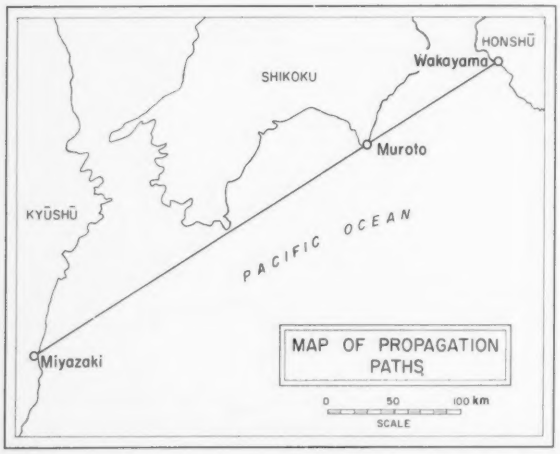


FIGURE 5. *Map of propagation paths.*

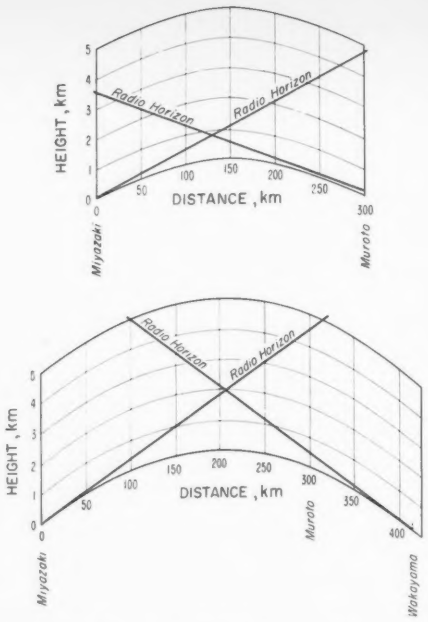


FIGURE 6. *Profiles of propagation paths.*

radiosonde observations were performed at Muroto four times a day in order to analyze the effects of radiometeorological conditions on radio propagation.

The 1,300- and 3,000-Mc hourly median basic transmission losses are used in the present analysis. Calculations of refractive index were made at each transition point of temperature code of radiosonde which covered the height interval from 0 to 5,000 or to 7,000 m.

In figure 7 is shown the variation of median basic transmission loss for both frequencies and for both propagation paths, compared with radiometeorological factors; they are: the gradient of refractive index in every 500 m classified into three grades as shown in the figure, and relative humidity at 1,000 and 3,000 m above sea level. It is seen in this figure that high signal strength is accompanied by steep N-gradient at a considerable height for both frequencies and for both propagation paths. These layers are mainly caused by an extremely dry airmass in a high pressure area lying above a humid airmass. It is shown by the relative humidity curves that elevated layers appear when a large difference of relative humidity occurs between 1,000 and 3,000 m. In this season of the year, high pressure areas periodically traverse Japan from west to east. In response to the traverse of high pressure areas, median transmission loss presents a similar variation with a period of about 4 days. Systematic diurnal variation of transmission loss was not observed at all. Variations seem to be dependent largely on the property of the airmass covering the propagation path.

Correlation coefficients were calculated between transmission loss and N-gradient for two paths and

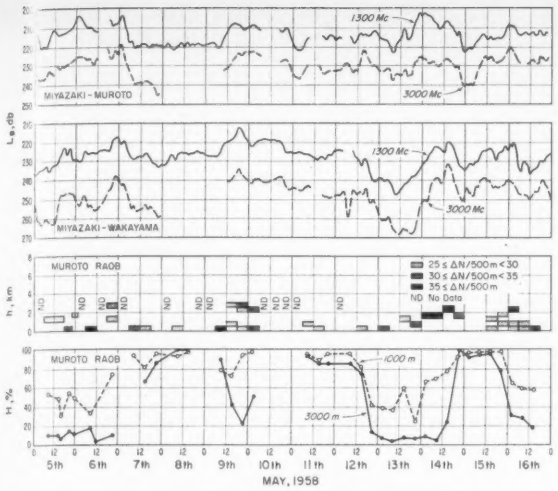


FIGURE 7. *Comparison of radio and meteorological data.*
For this figure only, *H* refers to the relative humidity in percent.

for the frequency of 3,000 Mc. In this case, the maximum value of $\Delta N/500$ m in the height range higher than a height H was correlated with transmission loss, considering that all layers located higher than H are concerned with the received field strength. Correlation coefficients calculated for various values of H are plotted against H in figure 8. It is interesting to note that the maximum value of the correlation coefficient is observed near the crossover height of each propagation path, and that

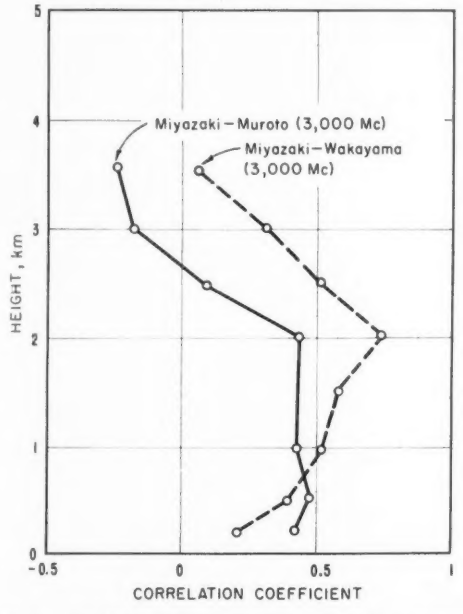


FIGURE 8. *Variation of correlation coefficient between median path loss and* $(\Delta N/500\text{ m})_{max}$ *with height.*

larger values of correlation coefficient are obtained on the Miyazaki—Wakayama path, the crossover point of which is closest to the radiosonde observation point.

No significant correlation was found, on the other hand, between transmission loss and N-gradients in the lowest 50, 200, and 1,000 m of the atmosphere.

A comparison with surface weather maps showed that high field strengths and elevated layers both accompanied high pressure areas, as stated above. A typical case, observed between the 12th and 15th of May when a high pressure area passed over the propagation paths, was analyzed in order to present the effect of a layer in more detail. Figure 9 shows a time cross section of relative humidity up to 7,000 m above sea level compared with variations in transmission loss during the same period. Equi-relative humidity contour lines are drawn in the figure for every 10 percent; an area of closely spaced lines denotes the existence of a layer with steep N-gradient. An extremely dry airmass, with relative humidity less than 10 percent, perhaps due to subsidence in the rear of the high pressure area, appeared on the afternoon of the 12th of May, producing a marked elevated discontinuity layer between 1,000 and 2,500

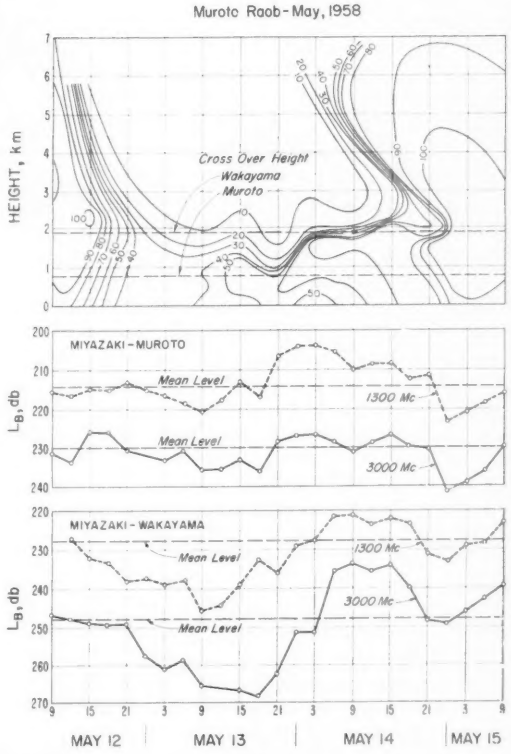


FIGURE 9. *The cross section of relative humidity in a high pressure area compared with transmission loss variation.*

m by the afternoon of the 14th. Comparing the transmission loss curves, shown in the lower part of the figure, with the variation of this layer, it is seen that high field strength started on the Miyazaki—Muroto path coincidently with the layer appearance at 21^h on the 13th and ended when it disappeared at 21^h on the 14th. Further, high signal strength started on the Miyazaki—Wakayama path when the layer reached up to the height of crossover point of the path at 06^h on the 14th. This shows that an appearance of an elevated layer higher than the crossover height produces a high field strength at the receiving point. It is also interesting to note that no change was found in the distribution of short term fading, which was represented by a Rayleigh distribution, during the whole process from the lowest through the highest field strength. This fact indicates that even a marked elevated layer does not act as a specular reflection plane, but produces many rays, numbering three or more, by finer structure of the layer.

As far as this propagation test is concerned the effects of the elevated layer played the most important role in the variation of median transmission loss within the month, occasional discrepancies being seen between them, perhaps due to other meteorological factors.

3.2. Experimental Results (NBS)

This section describes an attempt to determine relations between transmission loss and the presence of elevated layers, using experimental data obtained by NBS. Radiosonde data observed at Denver four times a day are used. By comparison, it was found that the effect of an elevated layer was not apparently distinguishable during this test period, because it was masked by more remarkable diurnal variations in the lower atmosphere. Two typical cases in August 1952 are shown in figures 10 and 11, when extreme enhancements in field strength were observed far beyond the horizon. In the first case (fig. 10), a marked discontinuity layer lies near the crossover height of the Pikes Peak—Anthony path. However, it is difficult to ascertain the effect of this layer because of a superrefractive layer observed simultaneously in the lowest part of the atmosphere. The fact that a similar enhancement occurs on the Pikes Peak—Fayetteville path, the crossover height of which is as high as 12,000 m above the ground, may emphasize the importance of superrefraction.

In the second case (fig. 11), we may be forced to attribute the enhancement occurring on the Camp Carson—Garden City path at 12^h to 14^h on the 11th to the layer observed near the height of crossover point (about 1,700 m), while we find a similar difficulty as in the first case (fig. 10) for the enhancements observed at Garden City and Anthony in the morning on the 12th. Further investigations would be necessary with careful treatments of lower superrefractive layers.

3.3. Considerations on Contribution of Layers of Propagation Far Beyond the Horizon

In the experiments by ECL, we found that elevated layers observed by radiosonde were the main factor in propagation far beyond the horizon. The lowest field level, which is expected in a linear or a tapered atmosphere, might be determined by scattering, internal reflection, or diffraction. However, this idealized type of the atmosphere was rarely observed in practice, the actual profile condition being characterized by sudden changes in the first derivative of N-profile. We could expect more and finer changes in the atmosphere that were missed by radiosonde because of its insufficient accuracy. Field strength far beyond the horizon seems to be largely contributed by laminar structures in the atmospheric refractive index, even though the lower or the lowest level might be determined by other modes of propagation.

In case of the experiments by NBS, on the other hand, we could not find so apparent effects as in the former case, mainly because of radiosonde data which were available only at significant levels and also be-

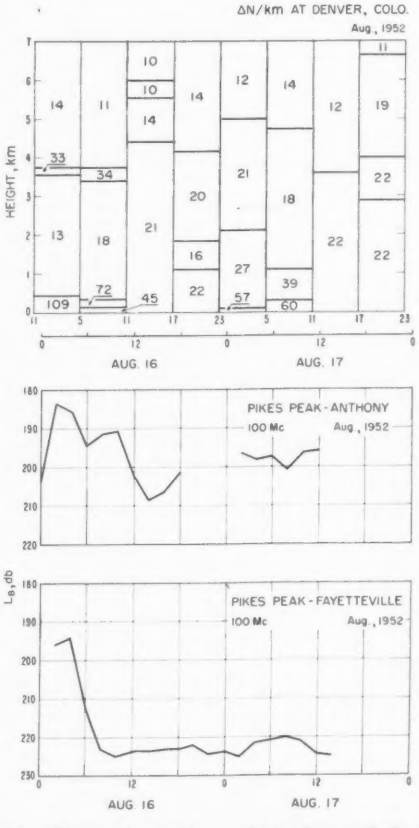


FIGURE 10. *Comparison of transmission loss with the gradient of refractive index of the atmosphere.*

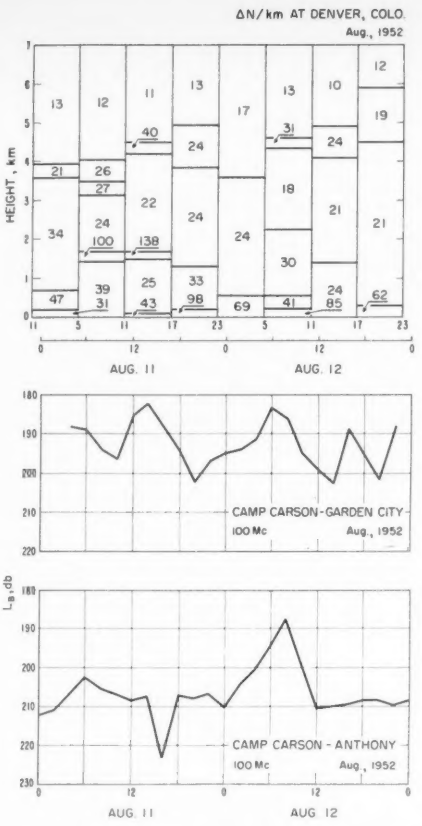


FIGURE 11. *Comparison of transmission loss with the gradient of refractive index of the atmosphere.*

cause of remarkable diurnal variations in the lower atmosphere. It is felt that the lowest atmosphere occasionally plays a role, probably by superrefraction, but is not frequent enough to determine the general characteristics of the variation in transmission loss. A study reported elsewhere [13] shows a similar result on the contribution of elevated layers and also of superrefractive layers, based upon experimental data obtained on an overland propagation path. Considerations of the lowest atmosphere should be made in further studies to distinguish its effects from those of the upper atmosphere.

4. Suggestions as to the Probable Importance of Radiometeorological Factors on Beyond-Horizon Communications

Preparing for the further studies on prediction of variations in transmission loss within a month, we tried to disclose dominant radiometeorological factors in propagation beyond the horizon. Near the horizon, where the diffraction mode is considered predominant, the gradient of refractive index at the earth's surface may largely affect the received field strength. However, effects of elevated layers seem to increase rapidly just beyond the horizon, emphasized by refraction in the lowest part of the atmosphere. In this region, diurnal variation in field strength may occur on an overland path corresponding to the variation in the atmosphere. Diurnal variation decreases as the crossover height increases, and diminishes at crossover heights higher than about 500 m above the ground. In propagation far beyond the horizon, no significant diurnal characteristic is found. At such ranges, the received field is mainly controlled by discontinuity layers that are dependent on airmass properties and may be occasionally affected by superrefraction.

This interpretation, which is not based upon a sufficient amount of data to give a final conclusion, may suggest, perhaps, that the effects of radiometeorological factors in the common volume play the most important role in communications beyond the horizon. It depends, of course, upon the property of propagation path, geographic situation, season, and climatological condition. Nevertheless, the atmospheric refractive index structure within the common volume looks like the most promising factor which determines the general characteristics of variations in transmission loss within a month, from the standpoint of practical communications. Perhaps a correction due to refraction may be necessary just beyond the horizon.

This interpretation would also be helpful for better understanding of the description of the seasonal variation in transmission loss in terms of surface refractivity or N-gradient at the earth's surface.

5. Conclusion

Propagation characteristics beyond the horizon should be discussed in two categories. In the first case, when crossover height is located in the diurnally varying portion of the atmosphere, diurnal variation in field strength should be taken into consideration. In the second case, when it is higher than about 500 m above the ground, field strength is largely dependent on the refractive index structure of the atmosphere free from the earth's surface. In both cases, the most dominant radiometeorological factor is considered to be the laminar structure of the atmosphere, which might be utilized for the expression of variations in transmission loss within a month.

The author thanks the United Nations which made possible his study in Boulder Laboratories, NBS, and the Federal Communication Commission, and National Bureau of Standards, Department of Commerce, without whose arrangements of his study program this paper could not have been written.

The author expresses his particular appreciation to B. R. Bean for his helpful advice and criticism as well as for his efforts in arranging radio and meteorological data obtained by NBS, and to H. B. Janes

for his valuable information and advice in the analysis of diurnal variation. He also thanks the many individuals in the Radio Propagation Engineering Division, Boulder Laboratories, NBS, and in the Radio Wave Propagation Section, ECL, N.T. & T., who contributed to the propagation experiments and to the reduction of the obtained data.

6. References

[1] B. R. Bean and F. M. Meaney, Some applications of the monthly median refractivity gradient in tropospheric propagation, Proc. IRE **43,** 1419 (1955).

[2] K. A. Norton, P. L. Rice, and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43,** 1489 (1955).

[3] R. E. Gray, The refractive index of the atmosphere as a factor in tropospheric propagation far beyond the horizon, IRE Natl. Conv. Record (1957).

[4] M. Onoe, M. Hirai, and S. Niwa, Results of experiment of long-distance overland propagation of ultrashort waves, J. Radio Research Labs., Japan **5,** 79 (1958).

[5] H. G. Booker and W. E. Gordon, A theory of radio scattering in the troposphere, Proc. IRE **38,** 401 (1950).

[6] T. J. Carroll and R. M. Ring, Propagation of short radio waves in a normally stratified troposphere, Proc. IRE **43,** 1384 (1955).

[7] O. Tukizi, Diffraction theory of tropospheric propagation near and beyond the radio horizon, IRE, Trans. AP–7, 261 (1959).

[8] H. T. Friis, A. B. Crawford, and D. C. Hogg, A reflection theory for propagation beyond the horizon, B.S.T.J. **36,** 627 (1957).

[9] H. B. Janes, An analysis of within-the-hour fading in 100-to-1,000-Mc transmissions, J. Research NBS **54,** 231 (1955).

[10] P. A. Sheppard, The structure and refractive index of the lower atmosphere, Meteorological factors in radiowave propagation, Conf. Rept. at Roy. Inst., London, England, p. 37 (1946).

[11] F. Ikegami, M. Haga, T. Fukada, and H. Yoshida, Observations of ducts in the lower atmosphere and their effects on microwave fading (in Japanese), Elec. Commun. Lab., NT & T Corp., Tokyo, Japan, rept. No. 910 (1956).

[12] T. Fukami, S. Ieiri, F. Ikegami, and H. Fujimura, Beyond-horizon propagation characteristics over sea (No. 1) (in Japanese) Elec. Commun. Lab. Tech. J., Tokyo, Japan, **8,** No. 6 (1959).

[13] B. R. Bean, Sur l' utilisation des observations météorologiques courantes en propagation radio électrique, L' Onde Elec. **362,** 411 (1957).

BOULDER, COLO. (Paper 64D3–53)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

The Trade-Wind Inversion as a Transoceanic Duct[1]

M. Katzin, H. Pezzner, B. Y.-C. Koo, J. V. Larson, and J. C. Katzin

(December 21, 1959)

Radiosonde data for stations in the South Atlantic trade-wind belt are analyzed to determine the potentialities of the trade-wind inversion as an elevated duct for transoceanic radio transmission. These were supplemented by refractometer soundings made by an aircraft during the latter part of 1958. These records indicate that a duct is present in the majority of the cases. Since it is known that the radiosonde underestimates ducting because of its slow response, it is concluded that a duct is present practically all the time. On the basis of the data analyzed, an experiment with two aircraft is suggested to test the propagation potentialities of this mechanism. A frequency of around 200 megacycles per second appears to be a good choice for an initial experiment.

1. Introduction

This paper describes the investigation of a conjecture that it may be possible to communicate across the oceans by means of vhf or uhf waves in the elevated ducts created by the trade-wind inversions. This conjecture developed out of the following evidence:

In World War II, record radar ranges were obtained on a 200-Mc radar at Bombay during the season of the northeast monsoon. Frequently the entire coast of Arabia would be painted on the radar PPI, and ranges up to about 1,700 miles were obtained. The meteorological situation, as reported by Durst [1],[2] results from hot dry air moving down from the Asian continent and overlying cooler moist air above the Arabian Sea. The combination of the two kinds of air produces a temperature inversion. At the level of the inversion, a temperature increase and a moisture decrease with increasing height occur. These both effect a decrease of refractive index with height, with the result that a strong elevated duct is created. Since a radar range of 1,700 miles represents a two-way traverse, the total range of transmission was up to 3,400 miles.

In a preliminary study of meteorological data for the South Atlantic region between Ascension Island and the Brazilian coast, a similarity of the conditions in the trade-wind inversion to those which were effective in producing the record ranges over the Arabian Sea became apparent. Since the trade inversion extends between the continents, or islands contiguous thereto, it appeared that here might be a mechanism capable of propagating radio waves clear across the oceans with low attenuation, and thus requiring relatively little power to effect communications. Because of the persistence of the trade winds, which blow virtually continuously through the year, is also appeared that this mechanism might be a reliable one.

In order to determine the possible success of an experimental program, it was decided to analyze existing radiosonde records, which were the only data then available for regions of interest.

It was realized from the outset that radiosonde data had great shortcomings for use in revealing the existence of ducts. This is due principally to the lag of the temperature and humidity elements in comparison to the rate of ascent of the ballon. It was felt, however, that the indications obtained from the radiosonde would be pessimistic, so that a favorable indication perhaps should be interpreted as a strong probability.

Among the various trade-wind regions, the region of the South Atlantic trade was singled out for study, especially the portion between Ascension Island and the Brazilian coast.

In this paper, the results of the analysis of available radiosonde data are summarized. Because of the favorable nature of these results, arrangements were made to take airborne refractometer measurements of the refractive index distribution through the inversion region. These measurements were made during a two-week period in November and December 1958 by Dr. R. M. Cunningham of the Air Force Cambridge Research Center in his Airborne Cloud Physics Laboratory. In addition, the regular daily schedule of radiosonde measurements at Fernando de Noronha and Ascension Island were stepped up to four per day during this period for comparison with the refractometer soundings, as well as to investigate a possible diurnal variation in inversion characteristics. The results of these additional measurements are also summarized here.

2. Characteristics of the Trade Inversion

A brief review of the known characteristics of the trade-wind inversion will be presented as a background against which to view the results obtained from the present study. Present knowledge of these characteristics is due largely to the classical papers of von Ficker [2] and Schnapauff [3], which are sum-

[1] Contribution from Electromagnetic Research Corp., Washington, D.C. This work was supported by Air Research and Development Command, U.S. Air Force.

[2] Figures in brackets indicate the literature references at the end of this paper.

marized in Riehl's [4] book. We shall be concerned here entirely with the Atlantic trade winds, especially those of the South Atlantic.

It is believed that the trade-wind inversion is due to dry air descending from high levels at the eastern ends of subtropical high-pressure regions. This subsiding air meets a current of moist and cooler air above the oceans. The warm air above forms a stable blanket over the cooler air below which inhibits vertical motion and large-scale mixing of the two streams of air. An inversion thus forms at the level where these streams meet.

According to von Ficker [2], the height of the trade inversion is subject to day-to-day and seasonal fluctuations. In the South Atlantic trade, the inversion rises as one proceeds westward from the African coast. The temperature rise, or "strength" of the inversion, decreases to the westward and equatorward. The most intense change in relative humidity within the inversion also is associated with the lowest inversion heights, and decreases westward and equatorward.

From the height, temperature rise, and humidity decrease characteristics, von Ficker's data indicate that the inversion begins at the African coast at a low height and is very intense. As one progresses westward, there is at first a relatively rapid rise in height, accompanied by a decrease in the amount of temperature and humidity change. These changes in height and strength of the inversion flatten off about one-third of the way across the ocean, after which the height remains fairly constant and the temperature and humidity changes decrease slowly with distance.

3. Raw Data and Method of Analysis

For the present study, the principal source of information was radiosonde data which, for the most part, consisted of tabulations of temperature and relative humidity at so-called "significant" and "mandatory" levels. From these values the modified index, M, was computed from

$$M=N+0.157h, \tag{1}$$

where h is the height in meters, and N, the refractivity, is calculated from the relation (Smith and Weintraub [5])

$$N=\frac{77.6}{T}\left(p+4{,}810\,\frac{e}{T}\right),$$

in which T is the absolute temperature, p is the atmospheric pressure in millibars, and e the partial pressure of water vapor in millibars.

An example of the results of a sounding is shown in figure 1. The curve (a) is an M-curve for a nighttime sounding taken at Ascension Island on November 7, 1944. From the M-curve, one deduces that an elevated duct with a ΔM, or "strength," of 8 units is present between the heights 1,400 to 1,690 m. This sounding was specifically selected to show the presence of a duct.

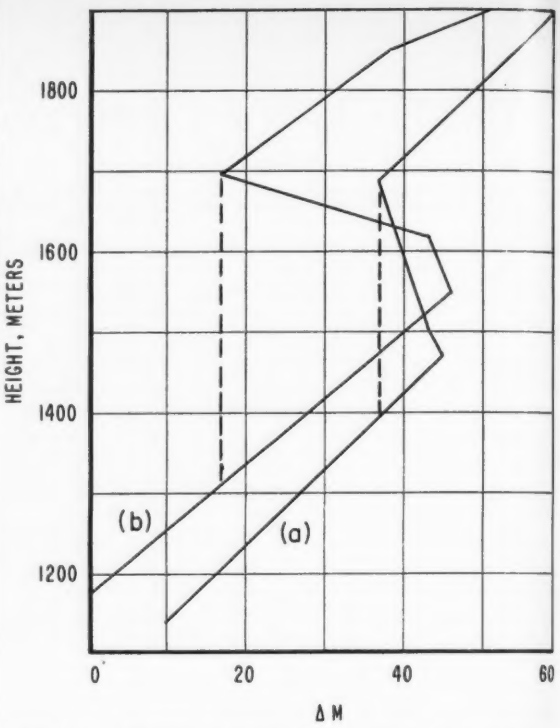


FIGURE 1. *M-curve calculated from radiosonde significant-level observations, Ascension Island.*

(a) Sounding of November 7, 1944, 2340 GCT; (b) sounding of November 15, 1957, 1145 GCT.

Because of the large amount of data, the following procedure was adopted for the bulk of the analysis. From the significant-level tabulations, the indicated levels of the bottom and top of the inversion could be spotted easily. From the values of pressure, temperature, and relative humidity corresponding to these and any intervening levels, N was computed. Chronological graphs were then prepared, showing the heights of the top and bottom of the inversion, the decrease of N between these heights, and surface wind speed and direction. Since some stations took one sounding per day, while others took two, the data were divided into day and night groupings for the latter cases. Some of these graphs are given in the accompanying figures, which will be discussed in section 4.

Dr. Cunningham's Airborne Cloud Physics Laboratory was equipped with extensive instrumentation for measurement of air and cloud particle properties. The data of immediate interest to the present study were the records of temperature and refractive index, especially the latter. The refractive index measurements were made with a microwave refractometer of the Crain type. A number of ancillary data were recorded, such as altitude, air speed, time markers, markers to indicate entry into and emergence from cloud, as well as cloud particle data. The profiles

through the inversion region were taken at a rate of ascent or descent of 500 to 1000 ft/min. Also a few horizontal runs were made near the cloud tops to explore the wavy nature of the base of the inversion.

4. Results and Discussion

4.1. Standard Radiosonde Summaries

The stations for which radiosonde summaries were analyzed are listed in table 1.

TABLE 1

Station	Location
Ascension I.	8.0° S, 14.6° W
Fernando de Noronha I.	3.8° S, 32.4° W
Luanda, Angola, P.W.A.	8.9° S, 13.0° E
Natal, Brazil	5.8° S, 35.3° W
Recife, Brazil	8.1° S, 34.9° W

Figure 2 is a chronological plot of the night soundings at Ascension Island for 1944. The points plotted for the individual days have been connected by straight lines. For some days data were missing, either due to no data or to no inversion. To differentiate between these two reasons for no data, dots were entered at the bottom of the graph, a single dot for no inversion and a double dot for no data. The instances of "no inversion" were quite rare.

A rather large day-to-day fluctuation in height and in ΔN is evident. In spite of this, a gradual seasonal trend in the heights can be discerned. Similar characteristics were found to exist in the daytime data, and in the data for 1945.

In order to bring out the seasonal characteristics clearly, 30-day running averages were computed for the heights and ΔN from the data of figure 2. These are plotted in figure 3. This brings out quite strikingly the seasonal variation, the average height of the lower boundary of the inversion increasing from about 1,300 m in summer (January) to about 1,800 m in winter (July). The thickness of the inversion shows only little variation between about 300 and 400 m, since the top and bottom move together fairly closely, even in detail. It must be remembered, however, that these are running averages of the data. In spite of this, the closeness with which the two curves follow each other is quite remarkable. From this, one is led to conclude that, by and large, whatever causes the day-to-day variations in height effects a bodily transport of the inversion layer. The average decrease of N in figure 3 varies between about 25 and 40 without, however, showing a marked seasonal correlation.

Data for the stations on or near the Brazilian coast varied considerably in quality. In most instances there are long gaps in the data, which make it difficult to deduce trends or averages. The data from the stations at Natal and Recife show evidences of several inversions existing at times.

As mentioned above in section 2, the sounding for curve (a) of figure 1 was specifically selected to show the presence of a duct. On that occasion a large decrease in N occurred over a moderate height interval, so that the first term on the right of eq (1) gave a larger decrease than the increase due to the second term containing h. Thus M decreased over the height interval between the two significant levels recorded. In the majority of cases, however, the

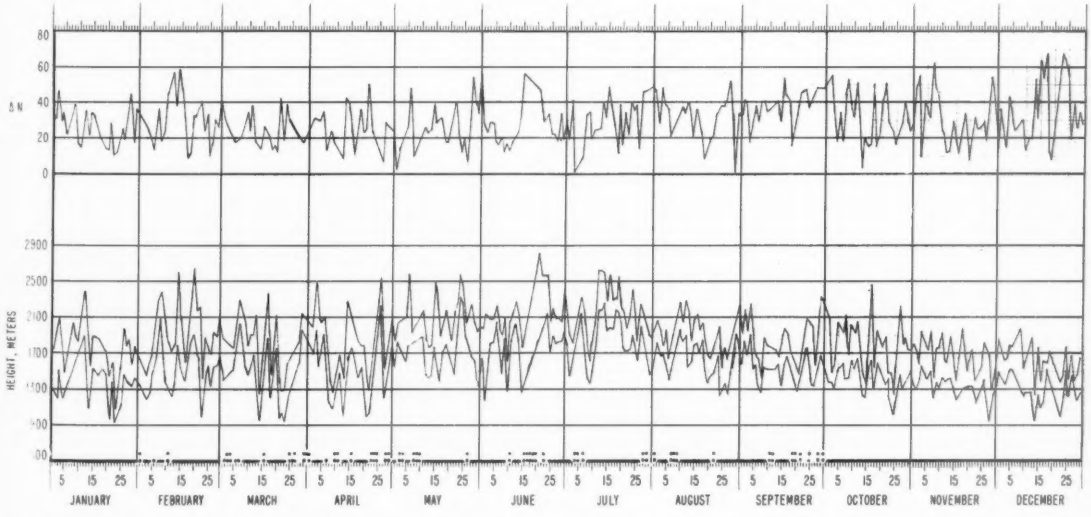


FIGURE 2. *Chronological plots of heights at top and bottom of inversion, decrease of N through inversion* (ΔN), *and surface wind speed and direction.*

Ascension Island, midnight soundings, 1944. Double dots at bottom denote "no data," single dots, "no inversion."

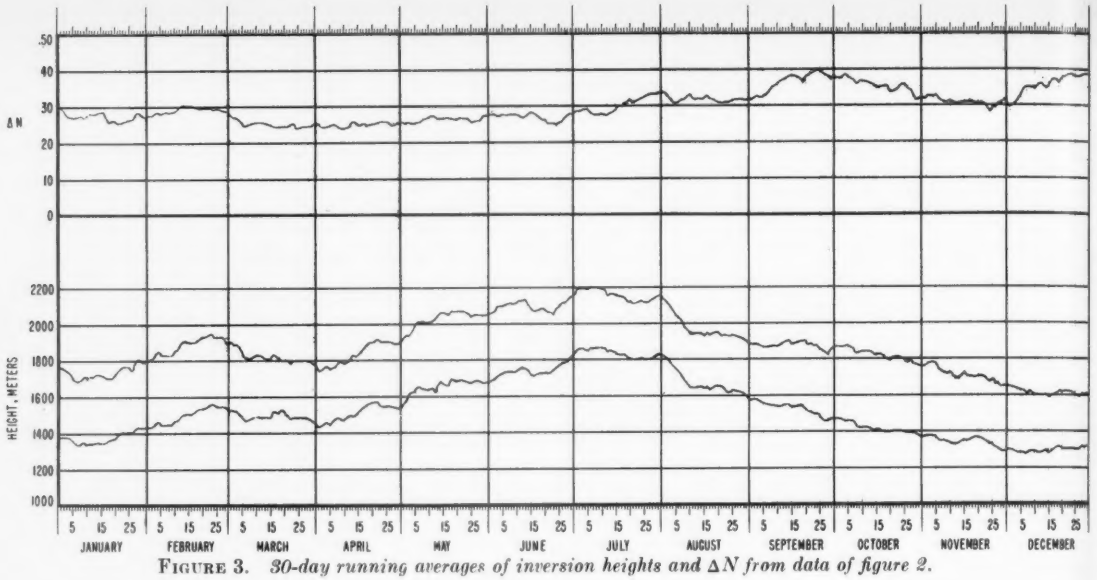


FIGURE 3. *30-day running averages of inversion heights and ΔN from data of figure 2.*

1944 and 1945 data from Ascension Island indicated no duct because the decrease in N took place over a height interval so large that the increase in the second term in eq (1) was larger than the decrease in N.

The data for Ascension Island for the latter part of 1957, shown in figure 4, are striking in comparison with the data in figure 2. It is evident immediately that the thickness of the inversion is much less in the later data, by a factor of two or more. Furthermore, the ΔN's run to higher values, also by a factor of about two. This has a profound effect on the trapping capabilities of the inversion. Curve (b) in figure 1 shows a plot of a selected sounding similar to that shown in curve (a). The M-curve shows a duct extending from 1,115 to 1,500 m, with a strength of 29 M-units. The same type of change is evident in the character of the later data for the Brazilian stations. The later data indicate that a duct is present in a majority of the soundings, as will be shown below.

The rather sharp change in inversion characteristics as revealed by the 1944 and 1957 data suggests that a change in the speed of response of the radiosonde instrument used took place between those dates (presumably the replacement of the hair hygrometer by the lithium chloride strip). This matter is related to the basic question regarding the reliability with which duct characteristics can be determined from radiosonde data. Andre [6] has discussed the limitations of radiosonde data for use in refractive index computations. He concluded that the modern radiosonde registers only those ducts which have scales of hundreds of feet or more, and that these are shown as being weaker, thicker, and higher than they actually are.

Table 2 shows the percentage of soundings which indicated ducts for the stations listed in table 1. It can be seen from table 2 that more recent measure-

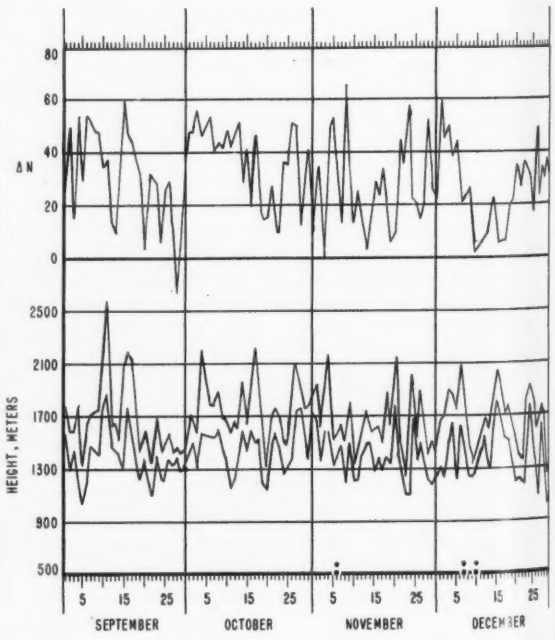


FIGURE 4. *Chronological plots of heights at top and bottom of inversion, and decrease of N through inversion (ΔN).*

Ascension Island, midnight soundings, 1957. Double dots at bottom denote "no data."

TABLE 2

Station	Period	Ducts
		%
Ascension I.	1943 to 1946	21
	1957 to 1958	60
Fernando de Noronha	1944	15
Luanda	1957 to 1958	1
Natal	1943 to 1946	14
Recife	1950 to 1952	8
	1954 to 1957	24

ment series show a much larger percentage of ducts present. (In the case of Luanda, which shows a very low percentage, it is known that this station still used the hair hygrometer.)

4.2. Four-Per-Day Radiosonde Measurements

The 4-per-day soundings at Fernando de Noronha and Ascension Island were scheduled to explore a possible diurnal trend in inversion height. The analysis of the significant level tabulations for inversion characteristics showed no significant diurnal variation, however.

During the two-week period, six soundings showed no inversion in the significant-level data. Inspection of the recorder charts, which were supplied for this period, showed that a large humidity lapse did occur through the normal levels of the inversion, in spite of no significant temperature change. Furthermore, examination of the charts showed that it was a consistent practice for the observers to smooth the records (in accordance with standard scaling procedures) in such a way as to yield a reduced humidity gradient. As a consequence, indications of the presence of a duct through analysis of significant level tabulations are made pessimistic, even beyond the effect of lag of the temperature and humidity sensing elements discussed earlier.

As a specific illustration, figure 5 shows plots of M versus height for the 1800 GCT sounding of November 30, 1958. At the time of this sounding, by a fortunate coincidence, a refractometer sounding also was taken in the same area by Dr. Cunningham. The dotted line is the plot of the significant level data, while the solid line is the plot of values scaled directly from the radiosonde recorder chart to reveal all the detail contained in the record. Also shown is the plot of the refractometer measurements made by Dr. Cunningham. Although the refractometer sounding shows two duct regions, the plot of the significant levels shows none, while the plot of the values scaled directly from the recorder chart indicates two ducts, each weaker and higher than the corresponding duct indicated by the refractometer plot. Thus this one comparison is in agreement with the general behavior predicted by Andre and mentioned in section 4.1.

4.3. Airborne Refractometer Measurements

In view of the rather large number of soundings taken by Dr. Cunningham, only a few samples of these will be included here. One already has been

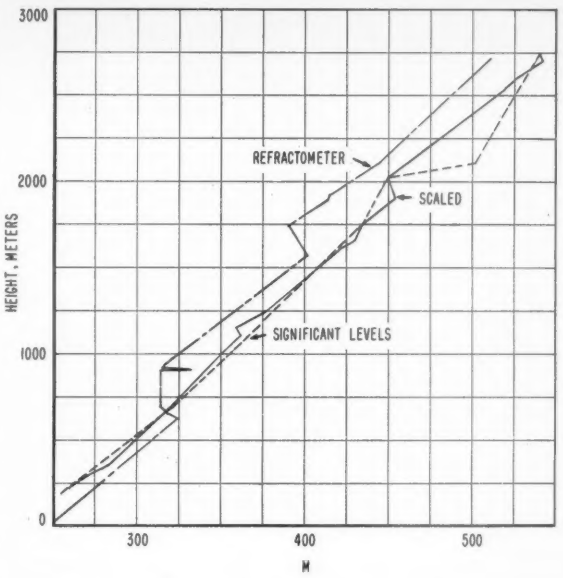


FIGURE 5. *Comparison of radiosonde and refractometer soundings.*

Fernando de Noronha, November 30, 1958.

shown in figure 5, and two others are shown in figure 6. The essential data relating to the duct region at the height of the trade-wind inversion for all of the refractometer soundings analyzed are given in table 3.

The soundings in figure 6 are typical of a number of cases in which the layer was indicated as having a serrated edge. This is due to the fact that the soundings actually are slant sections, however, rather than vertical profiles through the layer. The bottom of the inversion has a wavy profile, so that, at a vertical rate of the aircraft of 500 to 1,000 ft/min, the aircraft passes through several of the wave crests. This was verified in several horizontal runs through the cloud tops.

Table 3 lists the duct characteristics derived from the M-curves plotted for each profile. The table gives values of duct height, and thickness (in meters) and duct strength (ΔM) for each sounding, and the latitude and longitude corresponding to passage of the aircraft through the inversion. Of 68 soundings tabulated, only one showed no duct present. In two cases there was a gap in the recording because of a scale change made just at the time of passage through the inversion, so that values of duct thickness and strength could not be determined for these, although it was obvious that a duct existed.

Comparison of the aircraft soundings with the data derived from the radiosonde summaries shows that the inversion is much sharper and more intense than the radiosonde indicates. The transition frequently takes place in less than 30 m. As a result, stronger and thicker ducts are shown by the aircraft soundings.

TABLE 3

Date	Run	h	Δh	ΔM	°S	°W
11/30	2	1650	110	19	8.2	34.7
	3	1655	110	14	8.1	31.2
	4	1450	230	23	8.1	31.2
	5	1370	115	10	6.2	31.7
	6	1460	200	18	5.4	31.9
	7	1130	45	1	4.3	32.3
	8	1740	180	9	3.8	32.4
	9	1570	70	16	6.5	33.9
	10	1585	?	?	8.0	34.7
12/1	1	2040	?	34	8.1	34.7
	2	1540	165	11	8.5	29.5
	3	1495	130	8	8.3	30.6
	4a	1495	150	12	8.2	31.2
	4b	1515	245	15	8.1	32.1
	5	1470	280	18	8.0	33.0
	6	1610	200	13	8.0	34.1
	7	1510	260	44	8.2	34.5
12/2	2	1695	160	11	3.7	32.4
	3	2050	200	29	3.6	32.5
	4				6.1	33.9
	5	2050	335	23	6.7	34.2
	6	1670	90	6	8.0	34.8
12/5	1	1480	80	5	8.2	34.2
	2	1465	320	36	8.5	25.7
	3	1490	315	30	8.5	23.7
	4	1440	30	9	8.0	15.8
12/6	1	1305	160	17	8.2	14.2
	2	1445	245	19	7.6	20.1
	3	1500	235	17	7.7	19.4
	4	1405	365	31	8.0	17.7
	5	1285	170	16	8.1	17.0
	6	1450	235	22	8.1	16.1
	7	1485	105	10	8.1	15.5
	8	1440	235	27	8.7	14.9
12/7	1	1680	210	11	8.3	14.0
	2	1270	525	50	10.8	8.9
	4	1280	460	34	10.5	10.0
	11	1470	170	18	8.7	12.7
	12	1380	260	34	8.7	13.0
	13	1570	210	13	8.6	13.4
	14	1680	305	36	8.1	14.1
12/8	1	1300	175	21	8.0	14.3
	2	1410	410	30	8.2	20.3
	3	1380	400	46	8.2	19.8
	4	1435	410	27	8.2	19.5
	5	1375	150	15	8.1	19.1
	6	1430	580	45	8.1	18.7
	7	1390	>510	50	8.1	18.3
	8	1350	390	47	8.1	18.0
	9	1320	430	26	8.1	17.6
	10	1485	410	44	8.0	16.8
	11	1455	365	26	8.0	16.4
	12	1505	320	30	7.9	16.1
	13	1580	430	34	7.9	15.7
	14	1430	285	21	7.8	15.2
	15	1410	125	4	7.8	14.5
	16	1485	180	11	8.0	14.5
12/9	1	1595	105	11	8.0	14.5
	2	1300	405	34	7.9	25.3
	3	1500	430	32	8.0	26.0
	4	1550	255	42	8.2	26.7
	5	1410	110	4	8.3	27.5
	6	1720	140	10	8.2	34.5
12/10	10	1000	60	14	8.1	32.9
	11	1370	275	36	7.9	33.1
	12	1580	90	30	7.9	34.5
	13	1660	250	26	7.8	34.5
	14	1690	275	26	8.0	34.6

An approximate method of calculating the critical wavelength for trapping in an elevated duct composed of straight line segments was derived by the phase integral (WKB) method. When applied to the M-curves of figure 1, for example, the results shown in table 4 were obtained.

If it is assumed that the radiosonde underestimates the decrease in N and overestimates the height over which it takes place, then both the thickness of the duct and the magnitude of the decrease in N through it will be underestimated. Hence this will indicate

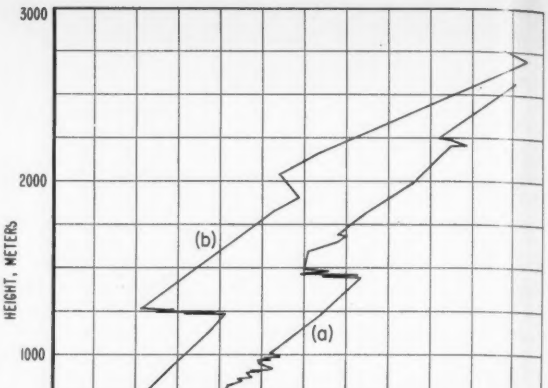


FIGURE 6. *M-curves from refractometer soundings.*

(a) Taken at 08.5° S, 25.7° W, December 5, 1958; (b) taken at 10.8° S, 08.9° W, December 7, 1958.

TABLE 4

	M-curve	
	Fig. 1	Fig. 9
λ_{max} (meters)	3.06	8.37
f_{min} (Mc)	98	37

a smaller trapped wavelength, and thus a higher minimum trapped frequency.

Out of the 68 soundings for which ΔM could be determined, it was found that only one would have a value of f_{min} above 200 Mc.

5. Conclusions

In view of the facts that

(*a*) the radiosonde consistently underestimates ducting conditions,

(*b*) the more recent data show ducts present in the majority of the soundings, and

(*c*) the aircraft soundings show consistently stronger ducts than the radiosonde data,

it is concluded that *a duct is probably present most, if not all, of the time.* From the observed duct characteristics, it is estimated that a conservative value of minimum trapped frequency is around 200 Mc.

An initial experimental test of the trapping capabilities of the trade wind inversion should aim at utilizing the potentialities in an optimum fashion. Thus *the waves should be launched in the duct and received in the duct.* Otherwise a large loss in exciting the trapped wave in the duct probably would be

incurred. From the data presented, it can be seen that the duct straddles the lower edge of the inversion. Hence a general guide is to *locate the transmitting and receiving antennas just below the base of the inversion.* This level generally is easy to determine because the cumulus clouds which are prevalent in the tropics generally terminate at the inversion base. The use of two aircraft for the transmitting and receiving terminals thus is indicated.

With regard to the choice of frequency, the effect of irregularities or "roughness" of the inversion boundaries and occasional "holes" in the inversion along the path must be considered. Roughness of the boundaries will result in a certain amount of scattering out of the duct, which should increase with frequency as the scattering irregularities become larger with respect to the wavelength. A "hole" in the duct will be somewhat equivalent to breaking a waveguide and separating the two parts. The transmitting end of the break will then radiate like a horn antenna and the receiving end will act like a receiving horn. The directivity of these equivalent horns will increase with frequency. Thus, if the axes of the horns are in line, a higher frequency should bridge the gap with lower attenuation than a lower frequency. But if the axes are out of line by an angle comparable to the beamwidth of the equivalent horn antenna, then a large loss will take place between the two ends of the break or hole. In this case a lower frequency would produce less loss. Because of the earth's curvature, an extensive hole in the duct would be equivalent to a misalinement of the axes at the apertures. From this it is concluded that frequencies not too far above the minimum trapped frequency should be used.

In making the final choice of frequency, other mechanisms of propagation which might interfere with or confuse the interpretation of the results should be considered. These mechanisms are meteoric and ionospheric scatter. The effectiveness of these mechanisms decreases rapidly with increasing frequency. At a frequency of 200 Mc these mechanisms should be feeble compared to the duct mechanism, so that such a frequency appears to be a good choice for an initial experiment.

6. References

[1] C. S. Durst, Radio climatology, Meteorological factors in radio-wave propagation, pp. 193–212 (Physical Society (London), 1946).

[2] H. v. Ficker, Die Passatinversion, Veroff, d. Met. Inst. d. Univ. Berlin, **1,** No. 4 (1936); translation into English by M. Flint, R.T.P. Translation No. 2371.

[3] W. Schnapauff, Untersuchungen Uber die Kalmenzone des Atlantischen Ozeans, Veroff d. Met. Inst. d. Univ. Berlin., **2,** No. 4 1937. English translation ASTIA Document No. A.T.I. 113162.

[4] H. Riehl, Tropical Meteorology (McGraw-Hill Book Co., Inc., New York, N.Y., 1954).

[5] E. K. Smith, Jr., and S. Weintraub, Constants in the equation for atmospheric refractive index at radio frequencies, J. Research NBS **50,** 39 (1953); Proc. IRE., **41,** 1,035 (1953).

[6] M. J. Andre, Limitations of radiosonde data for refractive index computations and note on services available from Air Weather Service, Proceedings of Symposium on Present and Future Uses of Refractive Index Data for Radio Propagation Purposes, Air Force Cambridge Research Center, (Sept. 6–7, 1956).

Boulder, Colo.

(Paper 64D3–54)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

An Analysis of Propagation Measurements Made at 418 Megacycles Per Second Well Beyond the Radio Horizon (a Digest)

H. B. Janes, J. C. Stroud, and M. T. Decker

(December 1, 1959)

During an 18-month period in 1952 and 1953, transmission loss measurements at 418 megacycles were made over a 134-mile path between Cedar Rapids, Iowa, and Quincy, Illinois. Continuous recordings made simultaneously at several receiving antenna heights from 30 to 665 feet yielded information on diurnal and seasonal variations in both the hourly median basic transmission loss and in height gain. These data are compared to predictions made using the method developed by Rice, Longley, and Norton and are found to be in good agreement, particularly at the lower antenna heights. An analysis of the correlation of short-term signal level variations observed at horizontally and vertically spaced antennas is described.

1. Introduction

This is a digest of an NBS Technical Note [1] [1] which describes in considerable detail a series of transmission loss measurements made by the National Bureau of Standards at a frequency of 418 Mc over a 134-mile path extending from Cedar Rapids, Iowa, to Quincy, Ill.

The principal purpose of the measurements was to study: (1) The hourly, diurnal, and seasonal variations in basic transmission loss experienced in transmissions made well beyond the radio horizon; (2) the corresponding long-term variability of height-gain; (3) the comparison of measured transmission loss and height-gain with predicted values; and, (4) the correlation of instantaneous signal levels measured at vertically and horizontally spaced antennas. (A discussion of item 4 is given in ref [1] but will not be summarized here.)

The experiment covered a period of approximately a year and a half from January 1952 to May 1953. The transmitter was located at Cedar Rapids and was operated by the Collins Radio Co. under contract with the National Bureau of Standards. The receiving and recording equipment were installed and operated by NBS. Space on a 750-ft tower was obtained through the cooperation of WTAD–FM in Quincy, and the receiving antennas were mounted on this tower at heights ranging from 30 to 655 ft above ground. The transmitting antenna was 39 ft above ground.

The equipment was operated for 13 recording periods, each of approximately 2 to 3 weeks duration. During each period, continuous recordings of basic transmission loss were made simultaneously at 3 to 5 different antenna heights. Table 2 of reference [1] shows a schedule of the recording periods and the antenna heights used during each period. Also shown are the inclusive dates during which each antenna height was used and the total number of hours of data recorded at each height during a given period.

The angular path distance [2] ranged from 20.3 to 15.7 milliradians for the 30- and 665-ft receiving antenna heights, respectively. Insofar as the long-term median basic transmission loss measured over paths having angular distances of this order (i.e., greater than about 10 milliradians) agree quite well with values predicted from scatter theory [2], this might be considered to be a tropospheric scatter propagation path. However, analysis of the short-term variations in signal level reveals that for significant percentages of the time (especially during the night), mechanisms other than scattering appear to be important. It would seem that this path is in a transitional region between the shorter paths where processes such as diffraction and ducting may provide most of the signal power and longer paths where scattering is the principal contributor.

2. Description of the Data

The data output of this experiment was reduced to hourly distributions of instantaneous signal levels obtained from time totalizing recorders or, in some instances, from paper chart recordings. The hourly median basic transmission loss, L_{bm}, and the fading range (ratio, in decibels, of levels exceeded 10% and 90% of the hour) were read from these distributions. A complete tabulation of all hourly L_{bm} values and fading ranges is given in reference [1]. Also given there are graphs of individual hourly L_{bm} values plotted versus time of day for each antenna height and each recording period. These graphs illustrate the fact that the range of the variations in L_{bm} measured at the same hour for 15 or 20 consecutive days is, in general, of the same order of magnitude as the range of the diurnal trend as shown by the median of hourly L_{bm} values.

Although the hourly fading ranges listed in reference [1] show considerable variance, they consistently cluster in the vicinity of the 13.4-db fading range that would be obtained from a pure Rayleigh distributed signal. This tendency is apparent regardless of time of day, although the variance of fading ranges measured at a given hour is somewhat larger

[1] Figures in brackets indicate the literature reference at the end of this paper.

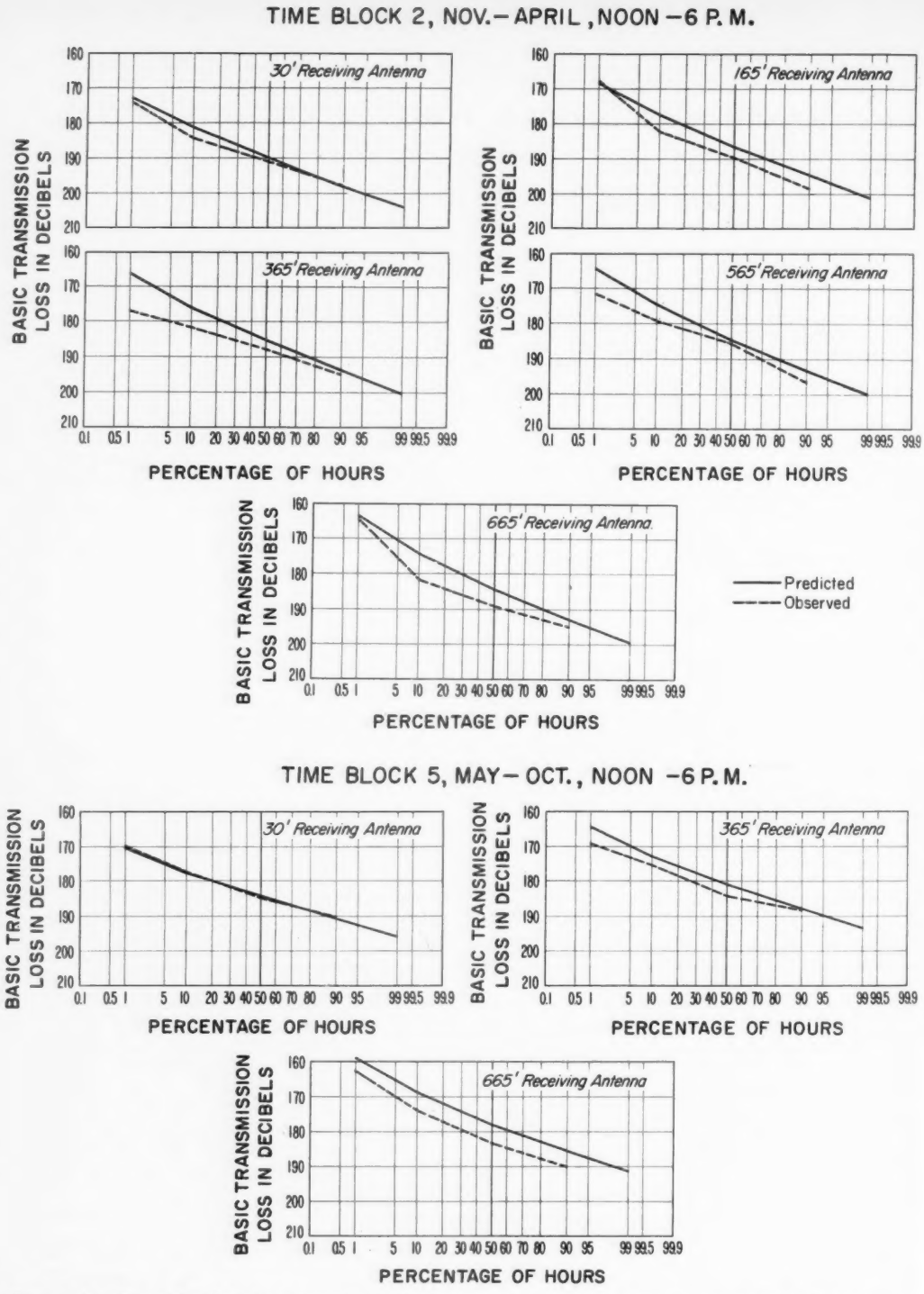

TIME BLOCK 2, NOV.– APRIL , NOON –6 P. M.
30' Receiving Antenna
165' Receiving Antenna
365' Receiving Antenna
565' Receiving Antenna
665' Receiving Antenna
BASIC TRANSMISSION LOSS IN DECIBELS
160 170 180 190 200 210
0.1 0.5 1 5 10 20 30 40 50 60 70 80 90 95 99 99.5 99.9
PERCENTAGE OF HOURS
Predicted
Observed
TIME BLOCK 5, MAY– OCT., NOON –6 P. M.
30' Receiving Antenna
365' Receiving Antenna
665' Receiving Antenna
BASIC TRANSMISSION LOSS IN DECIBELS
PERCENTAGE OF HOURS

FIGURE 1. *Observed and predicted cumulative distributions of hourly median basic transmission loss (1952 and 1953 data).*

during the night than during the day, particularly in the summer months. The lack of diurnal cycle in fading range and the tendency to approximate the Rayleigh distribution fading range should not, however, be interpreted as evidence that the signals recorded at Quincy were the result of scattering regardless of the time of day. It can be seen in the data samples shown in the technical note that the character of short-term variations in signal level changes considerably in going from afternoon to night. The slow, deep fading occurring at night often covered a range equal to, or greater than, the Rayleigh fading range.

One of the principal objectives of the Quincy experiment was to study the effects of antenna height on the signals received far below the radio horizon. During most of the recording periods simultaneous transmission loss recordings were made at three or more receiving antenna heights. We define the height-gain associated with two antenna heights to be the difference between hourly median L_b values measured simultaneously at these two heights. A positive height-gain indicates a reduction in L_{bm} with increasing height. The hourly height-gain values obtained for each pair of antennas during each recording period were plotted versus time of day. It was found that although the median of hourly L_{bm} values tended to decrease slowly with height, there was a large variance in the height-gain observed on any one pair of antennas at a given hour of the day during a given recording period. Examples of these height-gain versus time-of-day plots are shown in reference [1]. Also included are graphs showing the median of the hourly height-gain values for each recording period plotted versus the ratio of receiving antenna heights, and the median basic transmission loss, fading range, and height-gain versus time of year.

3. Comparison of Observed and Predicted Basic Transmission Loss

Rice, Longley, and Norton [3] have developed a method of predicting the cumulative distribution of basic transmission loss at frequencies above 10 Mc for wide ranges of path lengths, antenna heights, terrain configurations, and atmospheric refractive index gradients (the latter as deduced from observed surface values of refractivity). Using this method, the predicted cumulative distributions of basic transmission loss were determined for the 30, 165, 365, 565, and 665-ft receiving antenna heights for time block two (November through April, from noon to 6 p.m.) and for the 30, 365, and 665-ft antenna heights for time block five (May through October from noon to 6 p.m.). These antenna heights were chosen because of the relatively large amount of observed data available for comparison. The surface refractivity data used in this determination were the average of values for these time blocks during 1952 and 1953 obtained at the U.S. Weather Bureau stations at Des Moines, Iowa, and Joliet, Ill. These stations lie west and east of the propagation path, respectively. However, the data obtained at these two points are so well correlated that we may reasonably assume that they closely approximate conditions on the path.

The predicted distributions are shown in figure 1 along with the corresponding distributions of observed values. The latter include all hourly medians observed during the time block in both 1952 and 1953. Figure 2 provides a comparison of the observed height-gain data with the corresponding predicted values. A predicted value is shown at each ratio of antenna heights; the observed value for each ratio is simply the mean of the median height gain values for all recording periods. On the assumption that the observed decibel values of height-gain are normally distributed, there is a 68 percent probability that the true mean lies within the "wings" on the observed points. The fact that the predicted values are all above the measured values (and, indeed, lie outside the wings in most cases) indicates a consistent bias in the prediction which is larger at the higher ratios, i.e., those involving the 30-ft antennas.

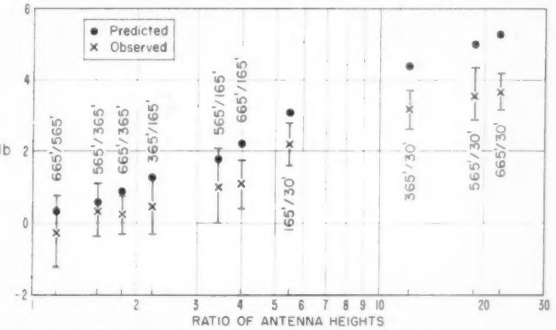


FIGURE 2. *Predicted and observed height-gain versus ratio of antenna heights for all hours 1952 and 1953.*

The experiment described here was performed under the supervision of J. W. Herbstreit. The surface refractivity data were supplied by B. R. Bean. Much of the reduction of the radio data was done by F. L. Anderson. K. A. Norton and P. L. Rice offered many valuable suggestions concerning the analysis.

4. References

[1] H. B. Janes, J. C. Stroud, and M. T. Decker, An analysis of propagation measurements made at 418 Mc well beyond the radio horizon, NBS Technical Note 6, (PB 151365) $2.25, available from the Office of Technical Services, U.S. Department of Commerce, Washington D.C. Foreign remittances must be in U.S. exchange and must include one-fourth of the publication price to cover mailing costs.

[2] K. A. Norton, P. L. Rice and L. E. Vogler, The use of angular distance in estimating transmission loss and fading range for propagation through a turbulent atmosphere over irregular terrain, Proc. IRE **43**, 10, 1488, (1955).

[3] P. L. Rice, A. G. Longley, and K. A. Norton, Prediction of the cumulative distribution with time of ground wave and tropospheric wave transmission loss, NBS Tech. Note 15 (PB151374) $1.50, available from the Office of Technical Services, U.S. Department of Commerce, Washington, D.C. Foreign remittances must be in U.S. exchange and must include one-fourth of the publication price to cover mailing costs.

BOULDER, COLO. (Paper 64D3–55)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

On the Calculation of the Departures of Radio Wave Bending From Normal

B. R. Bean and E. J. Dutton

(October 29, 1959; revised December 10, 1959)

The calculation of nonnormal tropospheric bending of radio waves is treated in terms of a reduced-to-sea-level value of the refractive index. This method emphasizes departures of bending from the average bending for the United States and consists of visualizing ray bending as consisting of two parts; an "average" component and a "departure-from-average" component. The "average" component comprises most of the bending and is obtained accurately from refraction tabulations while the component due to departures is easily obtained by graphical means.

1. Introduction and Background

The angular bending, τ, of a radio ray passing from the earth's surface where the refractive index is n_s to any point in the atmosphere of refractive index n is given by

$$\tau = -\int_{ns}^{n} \frac{\cot\theta}{n}\, dn \qquad (1)$$

where θ is the local elevation angle of the ray. (The geometry is shown on fig. 1.) In practice, (1) is evaluated by numerical integration, assuming that n is stratified in layers concentric with the earth and decreases linearly between measured or assigned values of n [1, 2].[1] Although at any instant the vertical distribution of n may be quite complex, it has been customary [3] to assume that n decreases on the average linearly with height up to ½ km, thus vastly simplifying many practical problems encountered in radio engineering applications involving antennas at these heights. Even though the assumption of a linear decrease of n is adequate for the majority of applications, the advent of space communications has forced a closer consideration of actual atmospheric n structure [4, 5, 6] with the general conclusion that it is more realistically represented by an exponential decrease with height than by a linear decrease. Millington [7] and Norton [8, 9, 10] have given numerous examples of the use of these more realistic atmospheres in the calculation of transmission loss in both ionospheric and tropospheric propagation, while Wait [11] has given rigorous series formulas for the distance to the horizon in such an atmosphere.

In fact, the International Radio Consultative Committee has recently recommended for international use a basic reference atmosphere to describe the average refractive index structure of the atmosphere by

[1] Figures in brackets indicate the literature references at the end of this paper.

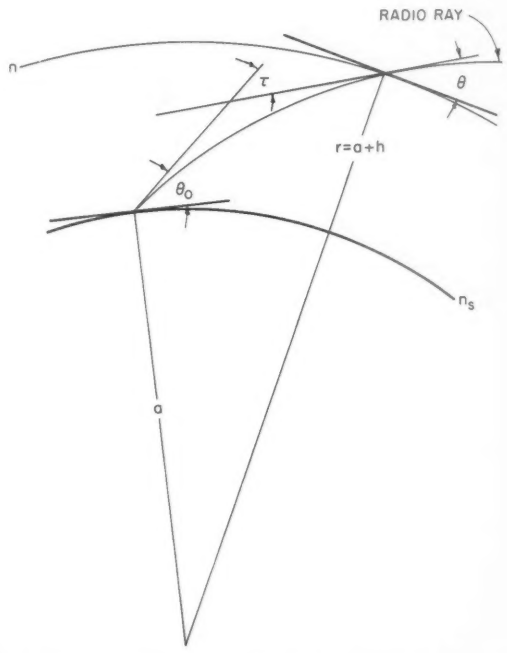


FIGURE 1. *Geometry of atmospheric refraction of radio waves.*

$$\overline{N(h)} = [\overline{n(h)} - 1]\, 10^6 = N_s \exp(-ch). \qquad (2)$$

Tables of refraction variables have been prepared [12] for atmospheres of this form for specific values of N_s and c throughout their normal range of variation.

The systematic exponential decrease of the refractive index with height is relatively large compared to variations introduced by fronts and airmass discontinuities. To obtain a clear picture of such variations one should eliminate the systematic height changes from the actual observed value,

THE ARRAY LIBRARY

$N(h)$, and thus emphasize the departures from normal of the actual N structure. This may be accomplished by

$$A(h, N_s)=N(h)+N_s\,[1-\exp(-ch)] \tag{3}$$

which has been found [13] to be more effective than previous corrections involving the assumption of a linear $N(h)$ distribution [14 through 17].

The purpose of this paper is to consider departures of n structure from normal, as specified by (2), the effect of these departures upon ray refraction, and to present a new method for calculating τ based upon these departures which will differ from other approximate methods already in the literature [1, 2, 5, 12, 18, 19, 20].

2. Calculation of Nonnormal Refraction

In evaluating the bending of a ray one is concerned with the refractive index gradient:

$$dA(h, N_s)=dN(h)+N_s c\,\exp(-ch)dh \tag{4}$$

or

$$10^6 dn\equiv dN(h)=dA(h, N_s)-N_s c\,\exp(-ch)dh. \tag{5}$$

The integral for the bending of radio waves,

$$\tau_{1,2}=-\int_{n_1}^{n_2}\frac{\cot\theta}{n}\,dn,$$

now becomes

$$\tau_{1,2}=-\int_{n_1}^{n_2}\frac{10^{-6}\cot\theta\,dA(h, N_s)}{n}-\int_{n_1}^{n_2}\frac{10^{-6}\cot\theta}{n}\,[-N_s c\exp(-ch)]\,dh. \tag{6}$$

The height variation of θ, although dependent upon the refractive index distribution through Snell's law, is at least four times more sensitive to purely geometric changes than to refractive index changes [2]. This then suggests that we may simplify our problem by assuming that θ may always be determined by its value in the $\overline{N(h)}=N_s\exp(-ch)$ atmosphere for the height under consideration. When this assumption is made the second term on the right-hand side of (6) becomes simply the bending in the $\overline{N(h)}=N_s\exp(-ch)$ atmosphere, designated $\tau(h, N_s)$. Equation (6) may now be written

$$\tau_{1,2}=-\int_{n_1}^{n_2}\frac{10^{-6}\cot\theta}{n}\,dA(h, N_s)+\tau\,(h, N_s). \tag{7}$$

Since $\tau(h, N_s)$ normally comprises most of the bending and is accurately tabulated (for several specific values of N_s), approximate or graphical methods may be used to determine

$$-\int_{n_1}^{n_2}\frac{10^{-6}\cot\theta}{n}\,dA(h, N_s). \tag{8}$$

The perturbation term represented by eq (8) presumably contributes only a small portion of τ which is taken as justification for the further approximations that $n=1$ and that $\cot\theta=1/\theta$, then

$$-\int_{n_1}^{n_2}\frac{10^{-6}\cot\theta}{n}\,dA(h, N_s)\approx-\int_{n_1}^{n_2}\frac{10^{-6}}{\theta}\,dA(h, N_s)$$

$$\approx\frac{-10^{-6}}{\bar{\theta}}\,[A(h, N_s)]_{n_1}^{n_2}=-\frac{2}{\theta_1+\theta_2}\,[A(h, N_s)]_{n_1}^{n_2}\,10^{-6}. \tag{9}$$

The problem now consists of determining the best method of estimating $\tau(h, N_s)$ and the variation of θ with height. One method would be to refer the method to a single exponential N profile typical of the United States:

$$\overline{N(h)}=313\,\exp\,(-0.144h) \tag{10}$$

where h is in kilometers. Equation (10) represents average conditions and would have the advantage of referring all N structure to a single "standard" atmosphere. However, climatic and synoptic changes in refractive index profiles can be accounted for by utilizing the range of refraction variables tabulated in reference [12]. These tabulations are in terms of specific values of N_s ranging from $N_s=200$ (extremely dry and high locations) to $N_s=450$ (extremely hot and humid locations). However, refraction effects are more sensitive to initial N *gradient* conditions than to the value of N_s [5], particularly at small elevation angles. Thus, in what follows, the choice of a reference atmosphere will be determined by the initial gradient of the N profile for which we wish to determine the bending. Table 1 lists the exponential profiles for which refraction tabulations are available. The N_s value

TABLE 1. *Initial* N *gradient*., ΔN_s, *in the C.R.P.L. exponential reference atmosphere* •

Range of ΔN_s (N units/km)	$N(h)$
$-\Delta N_s\leq 27.55$	$200\ \exp(-0.1184h)$
$27.55<-\Delta N_s\leq 35.33$	$252.9\ \exp(-.1262h)$
$35.33<-\Delta N_s\leq 42.13$	$289\ \exp(-.1357h)$
$42.13<-\Delta N_s\leq 49.52$	$313\ \exp(-.1438h)$
$49.52<-\Delta N_s\leq 59.68$	$344.5\ \exp(-.1568h)$
$59.68<-\Delta N_s\leq 71.10$	$377.2\ \exp(-.1732h)$
$71.10<-\Delta N_s\leq 88.65$	$404.9\ \exp(-.1898h)$
$88.65<-\Delta N_s$	$450\ \exp(-.2232h)$

• Note: height, h, is in kilometers.

and the range of initial N gradients assigned to each atmosphere are also listed. Since, theoretically, there are an infinite number of exponential atmospheres but, in practice, only a few with tabulated refraction variables, it will be necessary to assign each actual N profile encountered to a specific exponential reference atmosphere according to its initial N gradient. The initial N gradient of

$$\overline{N(h)}=N_s\exp\,(-ch)$$

is given by

$$\left(\frac{dN(h)}{dh}\right)_0 = -cN_s \tag{11}$$

For any actual case, however, one would not be able to define $(dN/dh)_0$ since the radiosonde reports at only certain points above the ground and not continuously. Thus if the first reporting level is at a height of ϵ, the actual initial gradient is determined from

$$-\Delta N_\epsilon = \frac{N_s - N_\epsilon}{\epsilon}. \tag{12}$$

For the practical application that is the objective of this paper we shall approximate $(dN/dh)_0$ by ΔN_ϵ, and the choice of an exponential atmosphere from table 1 will be based upon the observed value of ΔN_ϵ. As a consequence of this choice N_s will usually differ from $N(0)$.

The notation $A(h, 313)$ and $\tau(h, 313)$ indicates, for example, that the initial gradient of the particular profile under study falls into the 313 exponential atmosphere category, and that A, τ, and θ are determined from that particular atmosphere. Note that $A(0, 313)$ is not equal to 313 but $A(h, 313)$ will approach 313 at sufficiently large heights.

Radiosonde data from Santa Maria, California (a good test station because the refractive index profile departs quite radically from any smooth model) were graphically converted into $A(h, 313)$ units (see fig. 2) and the bendings calculated for several values of θ_0. The $A(h, 313)$ profile is compared on figure 3 with the departure of the bendings from the values expected in the 313 exponential atmosphere and emphasizes that τ is an integrated effect of all refractive index departures from the surface to the point under consideration while $A(h, 313)$ is determined solely from data at the height h alone. For instance, a strong surface-based refracting layer will bias all bending values above the layer while $A(h, 313)$ is independent of the values of N at any other height. It is quite

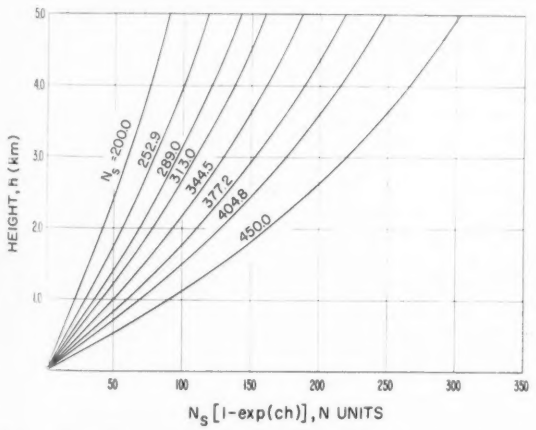


FIGURE 2. N_s *[1-exp(-ch)] versus height for conversion of* N(h) *profiles into* A(h, N_s) *profiles.*

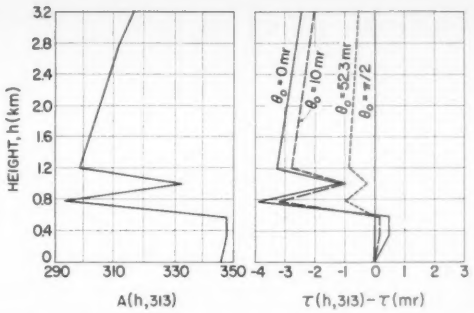


FIGURE 3. *Comparison of* A(h, *313) with the departures of ray bending from the value in the 313 exponential atmosphere for Santa Maria, Calif., Sept. 4, 1952, 1900 local time.*

Note that *negative* gradients of *A*(*h*, 313) produce *positive* values of ray bending.

clear that the agreement in height structure between the bending departures and $A(h, 313)$ is very pronounced on figure 3, especially in the layers nearest the ground and for small values of θ_0, indicating that a knowledge of $A(h, 313)$ alone could enable one to obtain a visual impression of the effect of nonnormal atmospheric structure upon the refraction of radio waves.

The above illustration has served to introduce a method of calculating τ in terms of departures of N structure from normal. The following section will be concerned with testing the accuracy of the method.

3. Comparison of Bendings

Significant level data for the radiosonde observations taken at 11 geographically and climatically diverse U.S. Weather Bureau observatories were analyzed to give adequate coverage of climatic conditions found within the cou try. Individual radiosonde observations taken during the years 1951 to 1952 were examined for each observing site, and profiles were selected to give a measure of the maximum variation of bending due to both geographic and local refractive index profile variations. The bending was then calculated for each of these profiles by both the standard numerical integration of (1), [1], and the departures-from-normal method described above. The differences between the bendings obtained by the two methods are given in table 2 as a percentage of the value obtained by conventional methods. A classification of the data of table 2 was made as regards height. Since most of the total bending to 70 km occurred within the first 3 km [5], the integral correction involving A was only applied to that height, the bending in the 313 atmosphere being used to estimate the average bending above 3 km thus assuming that the effect of N departures above this height is negligible. The mean errors are quite close to zero, indicating the absence of any marked bias in the method of calculation. One also notes from table 2 that the maximum percentage difference occurs at $\theta_0=0$ and generally decreases with increasing initial elevation angle.

TABLE 2. *Percentage deviation of ray bending values calculated by the departures-from-normal method from conventionally calculated values*

θ_0(mr)	0		10		52.3(3°)	
h(km)	3	70	3	70	3	70
	A. Bendings calculated relative to multiple exponential atmospheres					
Mean % of error	−0.84	−0.35	+0.49	+0.30	−0.007	−0.80
Std deviation of % errors	3.19	3.65	3.59	3.71	.97	3.95
Max % error	−10.0	+8.6	+9.4	+9.4	+2.5	+8.5
	B. Bendings calculated relative to the 313 exponential atmosphere					
Mean % of error	−0.03	+0.04	−3.04	−1.14	+0.78	−0.87
Std deviation of % errors	7.57	7.15	6.36	5.19	1.14	4.29
Max % error	−14.5	−14.0	−14.6	+10.9	−2.7	+9.0

Since the mean deviations are less than 1.0 percent the present system of calculation appears to be adequate for normal applications. One would probably not wish to use it for estimating τ for $\theta_0 > 52.3$ milliradians (mr) since at such high angles the total bending is as accurately estimated solely from ground level values by [18]:

$$\tau = N_s \times 10^{-6} \cot \theta_0. \tag{13}$$

One might also wish to use the departures-from-normal method referenced to a single average N profile typical of the entire United States. This approach was investigated by use of such an average, the 313 exponential atmosphere, and the results are tabulated in table 2. It is quite evident that the use of multiple reference profiles, classified according to ΔN_s, produces significantly smaller standard deviations and more consistent mean errors than does the use of a single average profile. This is, of course, what one would expect since the ΔN_s method adjusts the profile to the initial layer where a significant percentage of the total bending occurs. Interestingly enough, the maximum errors for $\theta_0 = 0$ arise from profiles such as shown on figure 2 where a very strong elevated layer exists well above the surface layer. Under this condition, the classification of profiles according to initial gradient does not account for the major anomaly of the profile. The maximum errors for $\theta_0 = 10$ mr are due to a quite different cause; the presence of ducting profiles in the test profiles. Under ducting conditions rays may be traced only for initial elevation angles greater than the angle of penetration [21] which, for radiosonde observed ducts, is less than 6 mr [22]. For this reason the data for $\theta_0 = 10$ mr includes many profiles that represent near ducting conditions and consequently maximum departures from normal refraction. In fact, the maximum errors for $\theta_0 = 10$ mr are associated with profiles having very thin surface ducts with normal or subnormal profiles above the duct, a condition that produces overestimation of the bending above the duct.

It should be observed that the $A(h, N_s)$ units as used in this paper are always referenced to zero *height* above ground level. We are able to ignore the *altitude* of the ground level by using multiple reference atmospheres which have, in effect, a built-in dependence of N_s upon ground level altitude. This effect may be accounted for with a single exponential atmosphere by adding the altitude of the ground level, b, to the height, h. Thus the equation for bending becomes

$$\tau = -\int_{n_1}^{n_2} \frac{10^{-6} \cot \theta}{n} \, dA(h, 313) + \exp(-cb)\tau(h, 313). \tag{14}$$

The exponential correction factor to $\tau(h, 313)$ is a constant for any particular station and is important only when b is large and consequently N_s appreciably less than 313.0.

One should bear in mind, however, that the above errors are about one-half that of other simple methods of estimating the bending of radio rays for $\theta_0 < 52.3$ mr such as regression techniques [18] or model atmospheres based upon the surface value of initial gradient conditions [5]. It would seem then that an additional advantage of the present system lies in the facility with which the user would come to appreciate the relative importance of various profile departures.

The data required for use of the present method are contained in reference [12]. Values of N_s [1-exp(-ch)] are given in table 3.

TABLE 3. *Values of* N_s [*1-exp*(-ch)]

Ht. (km)	200.0	252.9	289.0	313.0	344.5	377.2	404.9	450.0
0.01	0.2	0.3	0.4	0.4	0.5	0.6	0.8	1.0
.02	.5	.6	.8	.9	1.1	1.3	1.6	2.0
.05	1.2	1.6	2.0	2.2	2.7	3.2	3.9	5.0
.10	2.4	3.1	3.9	4.5	5.3	6.5	7.6	9.9
.20	4.7	6.3	7.7	8.9	10.6	12.8	15.1	19.7
.50	11.5	15.4	19.0	21.7	26.0	31.3	36.6	47.5
1.00	22.3	30.0	36.7	41.9	50.0	60.0	69.9	90.0
2.00	42.2	56.4	68.7	78.3	92.7	110.4	127.6	162.1
5.00	89.4	118.4	142.4	160.5	187.2	218.6	247.8	302.6

4. Conclusions

The present study has shown that deviations from normal ray bending are mirrored by departures of refractive index from values expected in average atmospheres of exponential form. Further, this departures-from-average method emphasizes that n-profile anomalies cause the greatest change in the bending when they occur near the ground, these changes being larger the smaller the initial elevation angle. A graphical method based upon these departures may be used to calculate the bending of radio rays.

The authors extend their gratitude to K. A. Norton and G. D. Thayer for their many critical and stimulating reviews of this work.

5. References

[1] M. Schulkin, Average radio-ray refraction in the lower atmosphere, Proc. IRE **40,** 554 (1952).

[2] L. J. Anderson, Tropospheric bending of radio waves, Trans. Am. Geophys. Union **39,** 208 (1958).

[3] J. C. Schelling, C. R. Burrows, and E. B. Ferrell, Ultra-short wave propagation, Proc. IRE **21,** 427 (1933).

[4] N. C. Gerson, Variations in the index of refraction of the atmosphere, Geofis. pura e appl. **13,** 88 (1948).

[5] B. R. Bean and G. D. Thayer, Models of the atmospheric radio refractive index, Proc. IRE **47,** 5, 740 (1959).

[6] P. Misme, Essai de radio climatologie d'altitude dans le nord de la France, Ann. des Telecommun. **13,** 11–12, 303 (1958).

[7] G. Millington, Propagation at great heights in the atmosphere, Marconi Rev. XXI, 143 (1958).

[8] K. A. Norton, Transmission loss in radio propagation—II, NBS Tech. Note 12 (PB151371) (June 1959). Office of Technical Services, U.S. Dept. of Commerce, Washington 25, D.C.; price $3.00.

[9] K. A. Norton, System loss in radio wave propagation, J. Research NBS **63D,** 53 (1959).

[10] K. A. Norton, Low and medium frequency radio propagation, Proc. Intern. Congr. Propagation of Radio Waves at Liege, Belgium, Oct. 1958, to be published by The Academic Press.

[11] J. R. Wait, Radio wave propagation in an inhomogeneous atmosphere, NBS Tech. Note 24 (PB151383) (Sept. 10, 1959). Office of Technical Services, U.S. Dept. of Commerce, Washington 25, D.C.; price $1.00.

[12] B. R. Bean and G. D. Thayer, The C.R.P.L. exponential reference atmosphere, NBS Mono. 4, Oct. 29, 1959. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C., price 45 cents. A digest of this monograph has appeared in the J. Research NBS **63D,** 315 (1959).

[13] B. R. Bean, J. D. Horn, and L. P. Riggs, Synoptic study of the vertical distribution of the radio refractive index J. Research NBS **63D,** 249 (1959).

[14] J. B. Smyth and L. G. Trolese, Propagation of radio waves in the lower troposphere, Proc. IRE **35,** 11, 1198 (1947).

[15] T. Y. Baker, On the refraction of electromagnetic waves in spherically stratified medium, Phil. Mag. **4,** 955 (1927).

[16] H. G. Booker and W. Walkinshaw, The mode theory of tropospheric refraction and its relation to wave-guides and diffraction. Meteorological factors in radio-wave propagation, p. 80, Rept. conf. Apr. 8, 1946 at Roy. Inst. London by Phys. Soc. and Roy. Meteorolog. Soc.

[17] Report of factual data from the Canterbury Project, Vols. I to III, Dept. Sci. Ind. Research (Wellington, New Zealand, 1951).

[18] B. R. Bean and B. A. Cahoon, The use of surface weather observations to predict the total atmospheric bending of radio waves at small elevation angles, Proc. IRE **45,** 11, 1545 (1957).

[19] G. H. Millman, Atmospheric effects on VHF and UHF propagation, Proc. IRE **46,** 1492 (1958).

[20] S. Weisbrod and L. J. Anderson, Simple methods for computing tropospheric and ionospheric refractive effects on radio waves, Proc. IRE **47,** 1770 (1959).

[21] D. E. Kerr, Propagation of short radio waves, Radiation Lab. Ser. **13,** p. 17 (McGraw-Hill Book Co., Inc., New York, N.Y., 1951).

[22] B. R. Bean, Climatology of ground-based radio ducts, J. Research NBS **63D,** 29 (1959).

Boulder, Colo.

(Paper 64D3–56)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

On the Mode Theory of Very-Low-Frequency Propagation in the Presence of a Transverse Magnetic Field

D. D. Crombie [1]

(December 14, 1959)

The effect of a purely transverse horizontal magnetic field on the propagation of very-low-frequency (vlf) waves is considered. It is shown that the magnetic field introduces nonreciprocity, and that for propagation along the magnetic equator, the rate of attenuation is less for west-to-east propagation than for east-to-west propagation.

1. Introduction

It has been shown by Budden [1][2] and by Wait [2] that the condition for vlf waves to be guided between a plane, perfectly conducting earth, and a plane, uniform, sharply bounded ionosphere is,

$$R=\exp\ (4\pi jhC_n/\lambda-2\pi jn) \tag{1}$$

where

n=order of mode,
h/λ=height of ionosphere in wavelengths,
R=the reflection coefficient of the ionosphere, and
C_n=the cosine of the complex angle of incidence θ_n of the wave on the ionosphere, for the nth mode.

It can be shown that the vertical component of the electric field of a given mode, n, at the ground in the earth-ionosphere waveguide is proportional to

$$\exp j\ (\omega t-2\pi dS_n/\lambda) \tag{2}$$

where S_n is the sine of the angle of incidence. If the wave is attenuated in the guide θ_n, C_n, and S_n are complex. A mode travels with a phase velocity

$$v_p=c/\mathrm{Re}\ S_n \tag{3}$$

and suffers an attenuation

$$A=kd\mathrm{Im}\ S_n,\ \text{nepers/unit distance.} \tag{4}$$

In these expressions, c is the velocity of light, d is the horizontal distance traveled, Re S_n and Im S_n are the real and imaginary parts of S_n, and k is $2\pi/\lambda$.

Budden [1] and Wait [2] have discussed the behavior of v_p and A for cases in which the earth's magnetic field can be neglected. Their treatment requires the substitution of (1) of the appropriate Fresnel reflection coefficient for a sharply bounded ionosphere, followed by the solution of (1) for C_n.

Barber and Crombie [3] have recently shown that if the ionosphere is assumed to be sharply bounded and the earth's magnetic field is taken to be purely transverse, then the reflection coefficient for waves polarized in the plane of incidence is given by the complex quantity

---

[1] Dominion Physical Laboratory, Department of Scientific and Industrial Research, Lower Hutt, New Zealand (invited paper).
[2] Figures in brackets indicate the literature references at the end of this paper.

$$R=\frac{C_n[(1+\delta)^2-\gamma^2]-[\mathrm{cm}_2(\delta+\delta^2-\gamma^2)/\omega-j\gamma S_n]}{C_n[(1+\delta)^2-\gamma^2]+[\mathrm{cm}_2(\delta+\delta^2-\gamma^2)/\omega-j\gamma S_n]}. \tag{5}$$

At vlf $\delta\cong j\beta=j\omega\nu/\omega_0^2$, $\gamma=\omega\cdot\omega_H/\omega_0^2$, where $\omega/2\pi$ is the wave frequency, $\omega_o/2\pi$ is the critical frequency of the ionosphere, and $\omega_H/2\pi$ is the gyrofrequency of the electrons in the ionosphere. The term cm_2/ω is given by

$$\mathrm{cm}_2/\omega=\sqrt{C_n^2+(1+j\beta)/(j\beta-\beta^2-\gamma^2)} \tag{5a}$$

the sign in front of the radical being chosen so that the imaginary part of cm_2/ω is negative. This ensures that the amplitude of the wave transmitted into the ionosphere decreases with height.

For propagation from west to east, the sign of γ is negative, while for east-to-west propagation it is positive. Thus the value of R, in general, differs for these two directions of propagation, and the earth-ionosphere waveguide exhibits nonreciprocal propagation as will be demonstrated below.

2. Propagation Constant of the Earth-Ionosphere Waveguide

Signals received over very great distances must suffer little attenuation; thus $|R|\cong 1$. Then, eq (5) can be written as

$$\left.\begin{aligned} R&=-(1-x)\\ \text{where } x&=\frac{2C_n[(1+j\beta)^2-\gamma^2]}{\mathrm{cm}_2(j\beta-\beta^2-\gamma^2)/\omega-j\gamma S_n}\end{aligned}\right\}. \tag{6}$$

From eq (1), taking natural logarithms

$$C_n=\frac{(n-1/2)}{2h/\lambda}+\frac{jx}{2kh}, \tag{7}$$

x is a function of C_n and S_n. Following Wait [2], a first-order perturbation solution of (7) can be obtained if C_n and S_n in x are replaced by $\overline{C}_n=\frac{(n-1/2)}{2h/\lambda}+$ and $\overline{S}_n=\sqrt{1-\overline{C}_n^2}$, the values obtained from (7) when $x=0$.

Then

$$S_n=\overline{S}_n-\frac{jx\overline{C}_n}{2kh\overline{S}_n}. \tag{8}$$

From (5a) and (6), since $\overline{C}_n{}^2$ is very small for low order modes,

$$x=-2\frac{\overline{C}_n[1-\beta^2-\gamma^2+2j\beta]\sqrt{j\beta}}{j\beta(1+j\beta)\sqrt{1-\gamma^2/j\beta(1+j\beta)}+j\gamma\overline{S}_n\sqrt{j\beta}}.$$

It is assumed that $\gamma^2<|\beta(1+j\beta)|$, and on taking $\sqrt{j}$ as $-(1+j)/\sqrt{2}$

$$x=\frac{2\overline{C}_n\sqrt{2\beta}(1-\beta^2-\gamma^2+2j\beta)(1+j)}{-2\beta^2-\gamma^2+\gamma\overline{S}_n\sqrt{2\beta}+j(2\beta-\sqrt{2\beta}\gamma\overline{S}_n}.$$

Separating this into its real and imaginary parts and substituting these in (8) gives

$$\mathrm{Im}\ S_n=-\frac{\overline{C}_n^2\sqrt{2\beta}}{kh\overline{S}_n}\cdot\frac{2\beta(1+\beta+\beta^2+\beta^3)+\gamma^2(3\beta^2+\gamma^2-1)-4\beta\sqrt{2\beta}\gamma\overline{S}_n}{4\beta\gamma^2\overline{S}_n^2+(2\beta^2+\gamma^2)^2+4\beta^2-2\gamma\overline{S}_n\sqrt{2\beta}(2\beta^2+2\beta+\gamma^2)} \tag{9}$$

$$\mathrm{Re}\ S_n=\overline{S}_n+\frac{\overline{C}_n^2\sqrt{2\beta}}{kh\overline{S}_n}\cdot\frac{2\gamma\overline{S}_n\sqrt{2\beta}(1-\beta^2-\gamma^2)-2\beta(1-\beta+\beta^2-\beta^3)-\gamma^2(1-3\beta^2-\gamma^2)}{4\beta\gamma^2\overline{S}_n^2+(2\beta^2+\gamma^2)^2+4\beta^2-2\gamma\overline{S}_n\sqrt{2\beta}(2\beta^2+2\beta+\gamma^2)}. \tag{10}$$

It is instructive to simplify these equations further by letting γ be much less than β. Equation (9) then reduces to

$$\text{Im } S_n = -\frac{\overline{C}_n^2}{2\sqrt{2}kh\overline{S}_n}\left(\frac{1}{\sqrt{\beta}}+\sqrt{\beta}\right)\left[1+\frac{\gamma\overline{S}_n\sqrt{2}}{(1+\beta)\sqrt{\beta}}\right]. \tag{11}$$

This is identical, when $\gamma=0$, to the value obtained by Wait [2] for propagation in the absence of the earth's magnetic field. Equation (11) shows that for west-to-east propagation, since γ is then negative, the attenuation coefficient is less than when the field is absent ($\gamma=0$) or when the propagation is from east to west (γ positive).

From eq (11) the ratio of the west-to-east, and east-to-west attenuation coefficients can be written as

$$\frac{\text{Im } S_n(W\to E)}{\text{Im } S_n(E\to W)} \cong 1-\frac{2\sqrt{2}\gamma\overline{S}_n}{(1+\beta)\sqrt{\beta}} \cong 1-\frac{2\sqrt{2}\omega_H\omega_\eta}{\nu^{3/2}\omega^{1/2}}$$

when β is greater than unity. The value of this ratio decreases as the wavelength increases, and thus the nonreciprocity becomes more marked at the lower frequencies.

When $\gamma \ll \beta$, eq (10) reduces to

$$\text{Re } S_n \cong \overline{S}_n + \frac{\overline{C}_n^2}{2\sqrt{2}kh\overline{S}_n}\left(\sqrt{\beta}-\frac{1}{\sqrt{\beta}}\right)$$

which again is the value obtained by Wait in the absence of a magnetic field. Thus the presence of a small transverse magnetic field does not affect the phase velocity of waves in the earth ionosphere waveguide under the conditions assumed here.

When the transverse magnetic field is such that γ and β are comparable, eqs (9) and (10) must be considered in full. Some calculations for likely ionospheric parameters have shown that nonreciprocity is still present under this condition.

3. Discussion

The writer has shown previously [4] that there is a considerable amount of evidence indicating that long nighttime vlf paths show nonreciprocal effects. The data refer to paths which cut the magnetic equator at various angles and thus the results cannot be really compared with the present work which deals only with a purely transverse horizontal, and small, magnetic field. It was not possible from the experimental evidence mentioned above to determine whether daytime paths exhibited reciprocity since in most cases the signals received over the daytime paths were too small.

However, when the ionosphere is dark reflection takes place from a greater height than when it is daylit, and thus the collision frequency ν is reduced considerably. Equation (11) shows that the effect of this is to make the nonreciprocity larger. Thus it may be expected that the effect of the transverse component of the earth's magnetic field will be smaller for a daytime path than for a nighttime path.

4. References

[1] K. G. Budden, The propagation of a radio atmospheric, Phil. Mag. **42,** 1 (1951).

[2] J. R. Wait, The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760 (1957). See also J. R. Wait, J. Research NBS **64D,** 153 (1960) for a discussion of the anisotropic case.

[3] N. F. Barber and D. D. Crombie, VLF reflections from the ionosphere in the presence of a transveres magnetic field (to be published in J. Atmospheric and Terrest. Phys.).

[4] D. D. Crombie, Differences between the east-west and west-east propagation of VLF signals over long distances, J. Atmospheric and Terrest. Phys. **12,** 110 (1958).

Boulder, Colo.

(Paper 64D3–57)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

On the Theory of Reflection of Low- and Very-Low-Radiofrequency Waves From the Ionosphere

J. Ralph Johler and Lillie C. Walters

(October 22, 1959)

The rigorous application of the magneto-ionic theory to the calculation of reflection coefficients for a sharply bounded ionosphere model is carried out. The paper is illustrated with computations applicable to the *D*-region or the *E*-region of the ionosphere. The quasi-longitudinal approximation is derived from this theory and the range of validity of this approximation is illustrated. The restrictions imposed by the use of a sharply bounded model ionosphere are discussed.

1. Introduction

The classical magneto-ionic theory introduced by Appleton [1],[1] Hartree [2,3], and Booker [4] describes the propagation of waves in a homogeneous, anisotropic ionosphere. Various methods have been proposed by Bremmer [5], Yabroff [6], and Budden [7] to calculate ionosphere reflection coefficients. The "quasi-longitudinal" (Q–L) approximation [7] provides considerable simplification of the calculation. Wait et al. [8] calculated vlf reflection coefficients with the aid of this mathematical approximation.

In this paper, the rigorous theory based on Maxwell's equations for a model ionosphere in which the electron density and collision frequency increase abruptly to a finite value at a particular altitude and so remain uniformly for all greater altitudes, is reduced to computational form. In this paper such a model is called a "sharply-bounded" model and represents a "first approximation" of the actual electron density-altitude, collision frequency-altitude profiles. The sharply-bounded ionosphere model is a physical approximation at low frequencies; the quality of the approximation decreasing with increasing frequency. The closeness of the approximation is also dependent upon the form of the ionosphere electron density-altitude profile near the assumed sharp boundary. The theory is illustrated by the calculation of certain reflection coefficients applicable to the *D*-region or the *E*-region of the ionosphere. Calculations are also presented to illustrate the range of validity of the Q–L approximation.

2. Theory

The xy-plane, figure 1, describes the boundary of the model ionosphere. The vertical direction, z describes the normal to the boundary. The region above the xy-plane ($z>0$) is characterized by a uniform electron density N and collision frequency ν. The region below the xy-plane ($z<0$) is a vacuum. A plane wave is incident on the boundary such that the normal to the wave front assumes an angle ϕ_i (angle of incidence) with the vertical z, and the plane of incidence, which can contain either electric $\overline{E}$ or magnetic $\overline{H}$ vector, is oriented at a magnetic azimuth ϕ_a i.e., the earth's magnetic field vector $\overline{H}_m$ is contained in the yz-plane, figure 1, with a magnetic inclination or dip I. The local coordinate system for the wave is described by x', y', z' such that the coordinates x, y, z and x', y', z' coincide when the angle of incidence ϕ_i and the magnetic azimuth angle ϕ_a vanish ($\phi_i=0$, $\phi_a=0$). An incident plane wave field $\overline{E_i}$ which varies harmonically in time t at a frequency $f=\omega/2\pi$ is incident upon the xy-plane,[2]

$$E_i=|\overline{E_i}|\exp\left\{i\left[\omega t-\frac{\omega}{c}\left(x\sin\phi_i\sin\phi_a+y\sin\phi_i\cos\phi_a+z\cos\phi_i\right]\right\}\right. . \tag{1}$$

[1] Figures in brackets indicate the literature references at the end of this paper.

[2] c=speed of light, $c\sim2.998\ (10^8)$ m/sec.

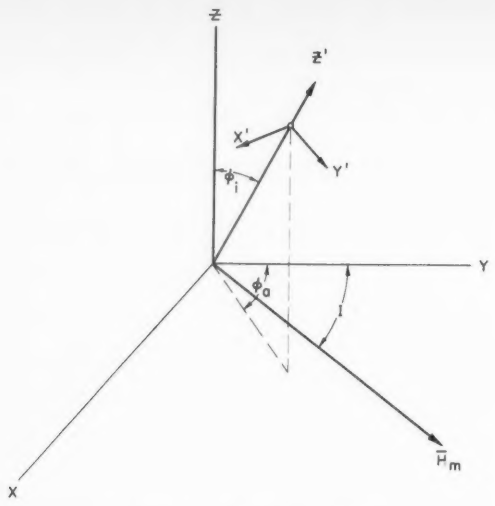


FIGURE 1. *Coordinate systems.*

A resultant wave transmitted into the ionosphere model ($z>0$) is then assumed to have the form,

$$E_t=|\overline{E}_t|\exp\left\{i\left[\omega t-\frac{\omega}{c}(x\sin\phi_i\sin\phi_a+y\sin\phi_i\cos\phi_a+z\zeta)\right]\right\}, \tag{2}$$

where ζ is, in general, a complex number. Such a wave is described by an electric $\overline{E}$ and a magnetic $\overline{H}$ intensity vector in Maxwell's classical equations,[3]

$$\bar{\nabla}\times\overline{E}+\mu_0\frac{\partial\overline{H}}{\partial t}=0, \tag{3}$$

$$\bar{\nabla}\times\overline{H}-\bar{J}-\epsilon_0\frac{\partial\overline{E}}{\partial t}=0, \tag{4}$$

where the permeability and dielectric constant of space are represented by μ_0 and ϵ_0, respectively, and where the electron convection currents are represented by the vector $\overline{J}=Ne\overline{V}$ for N electrons per cubic centimeter in which each electron has a charge e and travels at a vector velocity V; and the equation of motion of an electron,

$$m\frac{d\overline{V}}{dt}+m\nu\overline{V}+\mu_0 e(\overline{V}\times\overline{H}_m)+e\overline{E}=0, \tag{5}$$

where m is the mass of the electron and $\overline{H_m}$ is the earth's magnetic field intensity vector, describe the propagation in the ionosphere model. The simultaneous solution of these eqs (3), (4), (5) for a wave transmitted into the ionosphere $\overline{E_t}$ (2) upon elimination of the vectors $\overline{H}$ and $\overline{V}$ can be expressed as the matrix[4] equation,

$$\begin{bmatrix}A_x & A_y & A_z\\ B_x & B_y & B_z\\ C_x & C_y & C_z\end{bmatrix}\begin{bmatrix}E_x\\ E_y\\ E_z\end{bmatrix}=0. \tag{6}$$

$$A_x=1-a_L^2-\zeta^2-\frac{s}{s^2-h^2}, \tag{7}$$

[3] See for example, H. Bremmer [5] pp. 278 to 280.
[4] See for example, Yabroff [6].

$$A_y=a_La_T-i\frac{h_L}{s^2-h^2}, \tag{8}$$

$$A_z=a_T\zeta+i\frac{h_T}{s^2-h^2}, \tag{9}$$

$$B_x=a_La_T+i\frac{h_L}{s^2-h^2}, \tag{10}$$

$$B_y=1-\zeta^2-a_T^2-\frac{s^2-h_T^2}{s(s^2-h^2)}, \tag{11}$$

$$B_z=a_L\zeta+\frac{h_Lh_T}{s(s^2-h^2)}, \tag{12}$$

$$C_x=a_T\zeta-i\frac{h_T}{s^2-h^2}, \tag{13}$$

$$C_y=a_L\zeta+\frac{h_Lh_T}{s(s^2-h^2)}, \tag{14}$$

$$C_z=1-a^2-\frac{s^2-h_L^2}{s(s^2-h^2)}, \tag{15}$$

where,[5] defining the critical frequency squared $\omega_{cr}{}^2=Ne^2/\kappa m$ and the gyrofrequency $\omega_L=\mu_0H_m/m$,

$$s=\frac{\omega^2}{\omega_{cr}{}^2}\left[1-i\frac{\nu}{\omega}\right], \tag{16}$$

$$h=\frac{\omega_L\omega}{\omega_{cr}{}^2}, \tag{17}$$

$$h_L=-h\sin I, \tag{17a}$$

$$h_T=h\cos I, \tag{17b}$$

$$a_L=\sin\phi_i\cos\phi_a, \tag{18}$$

$$a_T=\sin\phi_i\sin\phi_a, \tag{19}$$

$$a=\sin\phi_i. \tag{20}$$

Rearrangement of the matrix eq (6) in powers of ζ results in a quartic equation in ζ with complex coefficients,[6]

$$a_4\zeta^4+a_3\zeta^3+a_2\zeta^2+a_1\zeta+a_0=0, \tag{21}$$

where,

$$a_0=(\sin^2\phi_i-1)^2\left[1-\frac{s}{s^2-h^2}\right]+(\sin^2\phi_i-1)\left[\frac{1}{s}+\frac{s-2}{s^2-h^2}+\frac{a_L^2h_T^2}{s(s^2-h^2)}\right]+\frac{s-1}{s(s^2-h^2)}, \tag{22}$$

$$a_1=2\frac{h_Lh_Ta_L}{s(s^2-h^2)}(\sin^2\phi_i-1), \tag{23}$$

$$a_2=\left\{2\left[1-\frac{s}{s^2-h^2}\right]+\frac{h_L^2}{s(s^2-h^2)}\right\}(\sin^2\phi_i-1)+\frac{h_T^2a_L^2}{s(s^2-h^2)}+\frac{s-2}{s^2-h^2}+\frac{1}{s}, \tag{24}$$

$$a_3=2\frac{h_Lh_Ta_L}{s(s^2-h^2)}=-a_1\sec^2\phi_i, \tag{25}$$

$$a_4=1-\frac{s^2-h_L^2}{s(s^2-h^2)}. \tag{26}$$

3. Determination of the Roots of the Quartic

The four complex roots of the quartic (21) were found by the Muller [9] iterative method applicable to n roots of the nth degree equation,

$$F(\zeta)=a_n\zeta^n+a_{n-1}\zeta^{n-1}+\ \ldots\ +a_0=0, \tag{27}$$

where a_n, a_{n-1}, $\ldots$ a_o are, in general, complex numbers and $a_n\neq0$. This method does not require the calculation of derivatives and necessitates only one value of the function per iteration.

[5] $\omega_{cr}{}^2=3.18\ (10^6)\ N$.
$\omega_L=1.76\ (10^7)\ H_m$, where $H_m=|\overline{H_m}|$ which by convention in geophysical data is called the earth's magnetic intensity in gauss (symbol: Γ) or the earth's magnetic intensity in gammas (symbol: γ) where $H_m=10^5\ \gamma.=\Gamma$.
$\kappa=1/c^2\ \mu_0=\epsilon_0$.

[6] This equation is equivalent to equations given by Bremmer [5] eq (18), p. 291 and Yabroff [6] p. 751.

Each iteration in the process of finding one root is obtained from a calculation of the nearer root of the quadratic equation which passes through the last three points of the function, $F(\zeta)$. This quadratic equation is, in general, complex, i.e., it possesses complex coefficients and complex roots. A LaGrange interpolation formula was employed in the process of finding the root. The function, $f(\zeta)$, in the LaGrange interpolation formulas,

$$f(\zeta)=A_n\zeta^n+A_{n-1}\zeta^{n-1}+\ \ldots\ +A_0, \tag{28}$$

is an nth degree equation with the same ordinates as $F(\zeta)$ for $n+1$ distinct abscissas [10] as $\zeta_n, \zeta_{n-1}\ \ldots\ \zeta_0$. The LaGrange interpolation formula through the three points; ζ_{i-2}, $F(\zeta_{i-2})$; ζ_{i-1}, $F(\zeta_{i-1})$; and ζ_i, $F(\zeta_i)$ of $F(\zeta)$ is,

$$f(\zeta)=A_2\zeta^2+A_1\zeta+A_0, \tag{29}$$

where A_2, A_1, and A_0 are evaluated from the requirement:

$$f(\zeta_{i-2})=F(\zeta_{i-2}),\ f(\zeta_{i-1})=F(\zeta_{i-1}),\ \text{and}\ f(\zeta_i)=F(\zeta_i).$$

Upon solving for A_2, A_1, A_0 and using the quantities $\beta=\zeta-\zeta_i$, $\beta_i=\zeta_i-\zeta_{i-1}$, $\beta_{i-1}=\zeta_{i-1}-\zeta_{i-2}$, $\lambda=\beta/\beta_i$, $\lambda_i=\beta_i/\beta_{i-1}$, and $\delta_i=1+\lambda_i$, the LaGrange interpolation formula, eq (29), becomes a quadratic in λ:

$$f(\zeta)=\lambda^2\delta_i^{-1}[F(\zeta_{i-2})\lambda_i^2-F(\zeta_{i-1})\lambda_i\delta_i+F(\zeta_i)\lambda_i]$$
$$+\lambda\delta^{-1}[F(\zeta_{i-2})\lambda_i^2-F(\zeta_{i-1})\delta_i^2+F(\zeta_i)\ (\lambda_i+\delta_i)]+F(\zeta_i). \tag{30}$$

ζ_{i+1} is found for the condition, $f(\zeta)=0$, solving for λ and employing the relationship,

$$\lambda=\lambda_{i+1}=\frac{\zeta_{i+1}-\zeta_i}{\zeta_i-\zeta_{i-1}}. \tag{31}$$

Rationalizing the numerator of the standard quadratic formula, and solving for λ_{i+1},

$$\lambda_{i+1}=\frac{-2F(\zeta_i)\,\delta_i}{g_i\pm\sqrt{g_i^2-4F(\zeta_i)\,\delta_i\lambda_i\,[F(\zeta_{i-2})\,\lambda_i-F(\zeta_{i-1})\,\delta_i+F(\zeta_i)]}}, \tag{32}$$

where,

$$g_i=F(\zeta_{i-2})\lambda_i^2-F(\zeta_{i-1})\delta_i^2+F(\zeta_i)(\lambda_i+\delta_i). \tag{33}$$

The sign choice in the denominator of λ_{i+1} (32) is resolved by selecting the value for the larger denominator. This choice of λ_{i+1} gives ζ_{i+1} the root of the LaGrange interpolation formula (29) which is nearer ζ_i. The initial values of ζ and $F(\zeta)$ can be taken as follows: $\zeta_0=-1$, $F(\zeta_o)=a_2-a_1+a_o$, $\zeta_1=1$, $F(\zeta_1)=a_2+a_1+a_0$, $\zeta_2=0$, and $F(\zeta_2)=a_0$. The iterative process is terminated with ζ_i, the derived root when,

$$\frac{|\zeta_i-\zeta_{i-1}|}{|\zeta_i|}<\epsilon,$$

where ϵ is a predetermined number.

Upon determination of a root, the degree of the polynomial can be reduced by dividing $F(\zeta)$ by the root. The new coefficients can be found by,

$$a_i'=\zeta_n a_{i-1}'+a_i,\ i=0,1,2\ \ldots,$$

with $a_{-1}'=0$, where a_i' replaces a_i and ζ_n is the root already calculated.

Two pairs of roots, ζ, where each root represents either an "ordinary" or an "extraordinary" wave propagated in the ionosphere, can be identified as upgoing ($+z$ direction, fig. 1, usually a fourth quadrant [7] ζ) and downgoing ($-z$ direction, fig. 1, usually a second quadrant ζ) waves. The upgoing pair determine the reflection coefficient for waves incident on the ionosphere from the region below the xy-plane ($z<0$).

[7] It is also possible to find roots in the 3d quadrant under certain conditions. See for example, Booker [4].

The quartic (21) is satisfied by the classical Appleton-Hartree formula,[8]

$$\eta^2=1-\frac{2\left(1-\frac{1}{s}\right)}{2\left(1-\frac{1}{s}\right)-\frac{h^2}{s^2}\sin^2\psi\pm\sqrt{\frac{h^4}{s^4}\sin^4\psi+\frac{4h^2}{s^2}\left(1-\frac{1}{s}\right)^2\cos^2\psi}}, \tag{34}$$

where,

$$\zeta^2=\eta^2-\sin^2\phi_i, \tag{35}$$

and,

$$\cos\psi=-\frac{\zeta}{\eta}\sin I+\frac{1}{\eta}\sin\phi_i\cos\phi_a\cos I, \tag{36}$$

$$\sin\psi=\frac{1}{\eta}\left[\eta^2-\zeta^2\sin^2 I-\sin\phi_i\cos\phi_a(\sin\phi_i\cos\phi_a\cos^2 I-2\zeta\sin I\cos I)\right]^{1/2}. \tag{37}$$

Although the Appleton-Hartree formula (34) provides a solution of the quartic for the upgoing waves (the dual sign in the denominator (34) represents ordinary and extraordinary waves), it is also necessary to satisfy the auxiliary eqs (35) to (37). But the iterative solution (27) to (33) is identical with such a process (34) to (37).

4. Determination of the Reflection Coefficients

The reflection coefficients are determined by five boundary conditions which express the principle of continuity of the tangential $\overline{E}$ and $\overline{H}$ fields and the normal $\overline{H}$ field [9] at the boundary (xy-plane, fig. 1) of the model ionosphere. The E_x field, figure 1, immediately above and immediately below the boundary can be equated,

$$E_{x'i}\cos\phi_a+E_{y'i}\cos\phi_i\sin\phi_a+E_{x'r}\cos\phi_a-E_{y'r}\cos\phi_i\sin\phi_a=Q_oE_{yo}+Q_eE_{ye}; \tag{38a}$$

also the E_y field,

$$-E_{x'i}\sin\phi_a+E_{y'i}\cos\phi_i\cos\phi_a-E_{x'r}\sin\phi_a-E_{y'r}\cos\phi_i\cos\phi_a=E_{yo}+E_{ye}; \tag{38b}$$

the H_x field,

$$E_{x'i}\cos\phi_i\sin\phi_a-E_{y'i}\cos\phi_a-E_{x'r}\cos\phi_i\sin\phi_a-E_{y'r}\cos\phi_a=(a_LP_o-\zeta_o)E_{yo}+(a_LP_e-\zeta_e)E_{ye}; \tag{38c}$$

the H_y field,

$$E_{x'i}\cos\phi_i\cos\phi_a+E_{y'i}\sin\phi_a-E_{x'r}\cos\phi_i\cos\phi_a+E_{y'r}\sin\phi_a=(\zeta_oQ_o-a_TP_o)E_{yo}+(\zeta_eQ_e-a_TP_e)E_{ye}; \tag{38d}$$

and the H_z field,

$$-E_{x'i}\sin\phi_i-E_{x'r}\sin\phi_i=(a_T-a_LQ_o)E_{yo}+(a_T-a_LQ_e)E_{ye}; \tag{38e}$$

where [10] (see fig. 1) $E_{x'i}$ is the electric field vector normal to the plane of incidence of the incident wave, $E_{y'i}$ is the electric field vector in the plane of incidence of the incident wave, $E_{x'r}$ is the electric field vector normal to the plane of incidence for the reflected wave, $E_{y'r}$ is the electric field vector in the plane of incidence for the reflected wave, also,

[8] See for example, H. Bremmer [5], p. 282, eq (8). The direction ψ is complex and double-valued (ψ_o, ψ_e) as a result of the two roots, ζ_o and ζ_e. The authors do not concur with a certain previous interpretation of the Appleton-Hartree formula, op. cit. [6], p. 751. It can be shown that the classical Appleton-Hartree formula, eqs (34) to (37), after considerable ado reduces precisely to the quartic, eq (21), the coefficients of which have been described, eqs (22) to (26); the latter of which are by definition valid for both zero ($\nu=0$) and finite ($\nu>0$) collision frequency. The authors must therefore conclude that the classical Appleton-Hartree formula is valid for finite collision frequency ($\nu>0$).

[9] See for example, Stratton [11], p. 483. The fifth boundary condition (38e) expresses the continuity of the normal $\overline{H}$ field. As a consequence of this condition (38e) it can be shown that (see eqs (41a–h)),

$$-T_{mm}-[(a_T-a_L\,Q_o)\,U_{mo}+(a_T-a_L\,Q_e)\,U_{me}]/\sin\phi_i-1=0.$$

The fifth condition (38e) must be satisfied automatically by the previous four conditions (38a–d), and hence the above expression (derived from 38e) can be employed as an independent partial check on the entire computation.

[10] The subscripts o and e refer somewhat arbitrarily to the two roots of ζ corresponding to "ordinary" and "extraordinary" waves.

$$Q=\frac{E_x}{E_y}, \tag{39}$$

$$P=\frac{E_z}{E_y}. \tag{40}$$

Four reflection coefficients and four "transmission" coefficients can now be defined,[11]

$$T_{ee}=\frac{E_{y'r}}{E_{y'i}}, \tag{41a}$$

$$U_{eo}=\frac{E_{yo}}{E_{y'i}}, \tag{41e}$$

$$T_{em}=\frac{E_{x'r}}{E_{y'i}}, \tag{41b}$$

$$U_{mo}=\frac{E_{yo}}{E_{x'i}}, \tag{41f}$$

$$T_{me}=\frac{E_{y'r}}{E_{x'i}}, \tag{41c}$$

$$U_{ee}=\frac{E_{ye}}{E_{y'i}}, \tag{41g}$$

$$T_{mm}=\frac{E_{x'r}}{E_{x'i}}, \tag{41d}$$

$$U_{me}=\frac{E_{ye}}{E_{x'i}}, \tag{41h}$$

where, T_{ee} refers to the plane wave ionosphere reflection coefficients with electric, $\overline{E}$, vector in the plane of incidence of the incident wave and the electric, $\overline{E}$, vector in the plane of incidence for the reflected wave; T_{em} refers to the same incident wave but the magnetic component (abnormal component) is in the plane of incidence for the reflected wave. Similar normal (electric) and abnormal (magnetic) components arise in the reflected wave resulting from an incident wave with magnetic component in the plane of incidence which components are described by the coefficients, T_{me} and T_{mm}, respectively. The quantities, U_{eo}, U_{mo}, U_{ee}, and U_{me} define the nature of an ordinary wave (second subscript o) and an extraordinary wave (second subscript e) excited in the model ionosphere by an incident wave with the electric vector in the plane of incidence (first subscript e) or magnetic vector in the plane of incidence (first subscript m) of the incident wave.

The matrix solution of the boundary conditions (38a, b, c, d) for an incident wave with the electric, $\overline{E}$, vector in plane of incidence,

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \begin{bmatrix} T_{ee} \\ T_{em} \\ U_{eo} \\ U_{ee} \end{bmatrix} + \begin{bmatrix} a_{oe} \\ b_{oe} \\ c_{oe} \\ d_{oe} \end{bmatrix} = 0, \tag{42}$$

and the matrix solution for an incident wave with the magnetic, $\overline{H}$, vector in the plane of incidence,

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \begin{bmatrix} T_{me} \\ T_{mm} \\ U_{mo} \\ U_{me} \end{bmatrix} + \begin{bmatrix} a_{om} \\ b_{om} \\ c_{om} \\ d_{om} \end{bmatrix} = 0, \tag{43}$$

completely define the four reflection coefficients: T_{ee}, T_{em}, T_{mm}, T_{me}; the "transmission" coefficients: U_{eo}, U_{ee}, U_{mo}, U_{me}, although of secondary interest, are a byproduct of the solution, where,

$$a_{oe}=\cos\phi_i \sin\phi_a, \tag{44a}$$

$$b_{om}=-\sin\phi_a, \tag{44d}$$

$$a_{om}=\cos\phi_a, \tag{44b}$$

$$c_{oe}=-\cos\phi_a, \tag{44e}$$

$$b_{oe}=\cos\phi_i \cos\phi_a, \tag{44c}$$

$$c_{om}=\cos\phi_i \sin\phi_a, \tag{44f}$$

[11] Bremmer's textbook notation for the reflection coefficient, T, is followed in this paper [5] pp. 286 to 295. Budden [7] and Wait [8] have employed for the reflection coefficients (41a, b, c, d) the notation $_{\parallel}R_{\parallel}$, $_{\parallel}R_{\perp}$, $_{\perp}R_{\parallel}$, $_{\perp}R_{\perp}$, respectively. However, for application to multihop geometric-optics (Bremmer [5] p. 194 et seq.) the "T" notation can be more readily discerned in the complicated formulas containing both ionosphere and ground reflection coefficients, in which the symbol R, has already been used for the latter.

$$d_{oe}=\sin\phi_a, \tag{44g}$$

$$d_{om}=\cos\phi_i\cos\phi_a, \tag{44h}$$

$$a_1=\cos\phi_a, \tag{44i}$$

$$a_2=-\cos\phi_i\sin\phi_a, \tag{44j}$$

$$a_3=-Q_o, \tag{44k}$$

$$a_4=-Q_e, \tag{44l}$$

$$b_1=-\sin\phi_a, \tag{44m}$$

$$b_2=-\cos\phi_i\cos\phi_a, \tag{44n}$$

$$b_3=-1, \tag{44o}$$

$$b_4=-1, \tag{44p}$$

$$c_1=-\cos\phi_i\sin\phi_a, \tag{44q}$$

$$c_2=-\cos\phi_a, \tag{44r}$$

$$c_3=\zeta_o-a_LP_o, \tag{44s}$$

$$c_4=\zeta_e-a_LP_e, \tag{44t}$$

$$d_1=-\cos\phi_i\cos\phi_a, \tag{44u}$$

$$d_2=\sin\phi_a, \tag{44v}$$

$$d_3=a_TP_o-\zeta_oQ_o, \tag{44w}$$

$$d_4=a_TP_e-\zeta_eQ_e. \tag{44x}$$

The solution of the matrices (42 and 43) was obtained by an application of Crout's [12] modified Gaussian method. Although it is quite possible to solve the matrix analytically and there are several numerical methods for the solution of matrices, the Crout method was employed because of its adaptability to electronic data processing and the requirement that the method chosen must be readily applicable to the solution of matrices with complex coefficients of the variables. The Crout method is applicable to n linear equations in n unknowns with either real or complex variables. The matrices (42) and (43) are of the form,

$$\begin{aligned}
&a_{11}x_1+a_{12}x_2+\ \ldots\ +a_{1n}x_n=c_1\\
&a_{21}x_1+a_{22}x_2+\ \ldots\ +a_{2n}x_n=c_2\\
&\ldots\ldots\ldots\ldots\ldots\\
&a_{n1}x_1+a_{n2}x_2+\ \ldots\ +a_{nn}=c_n,
\end{aligned} \tag{45}$$

where the n^2 coefficients a_{ij} and the n right-hand numbers c_i are given (44a-x). Since it is necessary to know the position of each element in the matrix to develop the method of solution, each element has a double subscript, that is, element a_{ij} is the element in the ith row and jth column.

The augmented or given matrix, $||a_{ij}||$ for the system (45) is formed by adjoining the right-hand numbers (the c's) to the matrix formed by the coefficients of the unknowns. This matrix can be expressed in rectangular array,

$$||a_{ij}||\equiv\left|\begin{array}{ccc|c}
a_{11}a_{12} & \ldots & a_{1n} & c_1\\
a_{21}a_{22} & \ldots & a_{2n} & c_2.\\
\ldots & \ldots & \ldots & \ldots\\
a_{n1}a_{n2} & \ldots & a_{nn} & c_n
\end{array}\right| \tag{46}$$

First the elements of the auxiliary matrix, $||a'_{ij}||$ are determined,

$$||a'_{ij}||\equiv\left|\begin{array}{ccc|c}
a'_{11}a'_{12} & \ldots & a'_{1n} & c'_1\\
a'_{21}a'_{22} & \ldots & a'_{2n} & c'_2.\\
\ldots & \ldots & \ldots & \ldots\\
a'_{n1}a'_{n2} & \ldots & a'_{nn} & c'_n
\end{array}\right| \tag{47}$$

From this matrix a final column matrix, $||x_i||$, with the elements for the required variables $x_1\ \ldots\ , x_n$ can be obtained,

$$\begin{matrix} x_1 \\ ||x_i|| \equiv x_2. \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{matrix} \tag{48}$$

Each element in both the auxiliary and final matrix is evaluated from previously calculated data by a sequential operation.

The following equations were used to calculate the elements a'_{ij} from the elements a_{ij} of the given matrix:

$$a'_{ij}=a_{ij}-\sum_{k=1}^{j-1} a'_{ik}a'_{kj},\ (i\geq j) \tag{49}$$

$$a'_{ij}=\frac{1}{a'_{ii}}\left[a_{ij}-\sum_{k=1}^{i-1} a'_{ik}a'_{kj}\right],\ (i<j) \tag{50}$$

$$c'_i=\frac{1}{a'_{ii}}\left[c_i-\sum_{k=1}^{i-1} a'_{ik}a'_{kj}\right], \tag{50a}$$

where any summation whose lower limit exceeds the upper is zero. Since each a'_{ij} in the auxiliary matrix, except those in the first row and first column, depends upon the value of previously calculated elements in the same matrix, the elements must be found in a certain order. The elements of the first column are found (49) as follows:

$$a_{11}=a'_{11},\ a_{21}=a'_{21},\ a_{31}=a'_{31},\ \text{and}\ a_{41}=a'_{41}.$$

The remaining elements of the first row are (50),

$$a'_{12}=\frac{a_{12}}{a'_{11}},\ a'_{13}=\frac{a_{13}}{a'_{11}},\ a'_{14}=\frac{a_{14}}{a'_{11}},\ \text{and}\ c'_1=\frac{c_1}{a'_{11}}.$$

With the aid of these values, the remaining elements can be found (49) or (50). a'_{ij} can be made available for the solution of the matrix $||a'_{ij}||$ by the following sequential calculation: The n elements of the first column, the remaining n of the $n+1$ elements of the first row, the remaining $n-1$ elements of the second column and second row, the remaining $n-2$ elements of the third column and third row, continuing in this sequence until the elements are calculated [9].

The values of x are found from x_n to x_1 from the elements of the auxiliary matrix,

$$x_i=c'_i-\sum_{k=i+1}^{n} a'_{ik}x_k. \tag{51}$$

It is evident (51) that $x_n=c'_n$. The remaining solutions $x_{n-1}, x_{n-2}\ \ldots\ , x_2, x_1$ can be found (51) with the aid of x_n. The computations presented in this paper were checked by substitution of $x_1, x_2, \ldots x_n$ in each of the original equations. An inductive-type proof of the Crout method has also been given by Crout [12] and Hildebrand [10].

The analytical expressions for the complex numbers P_o, P_e, Q_o, Q_e can be derived from the definitions (39) and (40) with the aid of the matrix (6) resulting from the simultaneous sc lution of Maxwell's eqs (3), (4), and the equation of motion of an electron (5) with the following result:

$$P=\frac{-\left[a_L\zeta+\frac{h_Th_L}{s(s^2-h^2)}\right]\left[1-a_L^2-\zeta^2-\frac{s}{s^2-h^2}\right]+\left[a_La_T-i\frac{h_L}{s^2-h^2}\right]\left[a_T\zeta-i\frac{h_T}{s^2-h^2}\right]}{\left[1-a^2-\frac{s^2-h_L^2}{s(s^2-h^2)}\right]\left[1-a_L^2-\zeta^2-\frac{s}{s^2-h^2}\right]-\left[a_T\zeta+i\frac{h_T}{s^2-h^2}\right]\left[a_T\zeta-i\frac{h_T}{s^2-h^2}\right]}, \tag{52}$$

$$Q=\frac{-\left[1-a^2-\frac{s^2-h_L^2}{s(s^2-h^2)}\right]\left[a_La_T-i\frac{h_L}{s^2-h^2}\right]+\left[a_T\zeta+i\frac{h_T}{s^2-h^2}\right]\left[a_L\zeta+\frac{h_Th_T}{s(s^2-h^2)}\right]}{\left[1-a^2-\frac{s^2-h_L^2}{s(s^2-h^2)}\right]\left[1-a_L^2-\zeta^2-\frac{s}{s^2-h^2}\right]-\left[a_T\zeta+i\frac{h_T}{s^2-h^2}\right]\left[a_T\zeta-i\frac{h_T}{s^2-h^2}\right]}, \tag{53}$$

where the subscripts o and e for P_o, P_e, Q_o, Q_e, refer to the two roots of ζ which represent upgoing waves in the model ionosphere.

5. The Q–L Approximation

Budden [7] and Wait [8] have evaluated reflection coefficients for the sharply bounded model ionosphere employing the mathematical Q–L approximation. This approximation has been employed to interpret experimental measurements of skywaves propagated at vlf between the surface of the earth and the ionosphere in the northern hemisphere.

Such an approximation can be derived from the rigorous formulation presented in this paper. The model ionosphere reflection coefficients can be written analytically from the rigorous matrix eqs (42) and (43),

$$T_{ee}=\frac{B_1A_2-A_1B_2}{A_3B_2-B_3A_2}, \tag{54}$$

$$T_{em}=\frac{B_1A_3-A_1B_3}{A_2B_3-B_2A_3}, \tag{55}$$

$$T_{me}=\frac{D_1A_2-C_1B_2}{A_3B_2-B_3A_2}, \tag{56}$$

$$T_{mm}=\frac{D_1A_3-C_1B_3}{A_2B_3-B_2A_3}, \tag{57}$$

$$A_1=\Delta_1^{-1}\left[\frac{a_{oe}}{a_4}-\frac{b_{oe}}{b_4}\right]-\Delta_2^{-1}\left[\frac{a_{oe}}{a_4}-\frac{c_{oe}}{c_4}\right], \tag{58}$$

$$A_2=\Delta_1^{-1}\left[\frac{a_1}{a_4}-\frac{b_1}{b_4}\right]-\Delta_2^{-1}\left[\frac{a_1}{a_4}-\frac{c_1}{c_4}\right], \tag{59}$$

$$A_3=\Delta_1^{-1}\left[\frac{a_2}{a_4}-\frac{b_2}{b_4}\right]-\Delta_2^{-1}\left[\frac{a_2}{a_4}-\frac{c_2}{c_4}\right], \tag{60}$$

$$B_1=\Delta_1^{-1}\left[\frac{a_{oe}}{a_4}-\frac{b_{oe}}{b_4}\right]-\Delta_3^{-1}\left[\frac{a_{oe}}{a_4}-\frac{d_{oe}}{d_4}\right], \tag{61}$$

$$B_2=\Delta_1^{-1}\left[\frac{a_1}{a_4}-\frac{b_1}{b_4}\right]-\Delta_3^{-1}\left[\frac{a_1}{a_4}-\frac{d_1}{d_4}\right], \tag{62}$$

$$B_3=\Delta_1^{-1}\left[\frac{a_2}{a_4}-\frac{b_2}{b_4}\right]-\Delta_3^{-1}\left[\frac{a_2}{a_4}-\frac{d_2}{d_4}\right], \tag{63}$$

$$C_1=\Delta_1^{-1}\left[\frac{a_{om}}{a_4}-\frac{b_{om}}{b_4}\right]-\Delta_2^{-1}\left[\frac{a_{om}}{a_4}-\frac{c_{om}}{c_4}\right], \tag{64}$$

$$D_1=\Delta_1^{-1}\left[\frac{a_{om}}{a_4}-\frac{b_{om}}{b_4}\right]-\Delta_3^{-1}\left[\frac{a_{om}}{a_4}-\frac{d_{om}}{d_4}\right], \tag{65}$$

where,

$$\Delta_1=\frac{a_3}{a_4}-\frac{b_3}{b_4}, \tag{66}$$

$$\Delta_2=\frac{a_3}{a_4}-\frac{c_3}{c_4}, \tag{67}$$

$$\Delta_3=\frac{a_3}{a_4}-\frac{d_3}{d_4}. \tag{68}$$

The Q-L approximations are as follows:

$$\phi_a=0, \tag{69}$$

$$Q_o=\frac{i}{\cos\theta_o}, \tag{70}$$

$$Q_e=\frac{-i}{\cos\theta_e}, \tag{71}$$

$$P_o=-\frac{\sin\phi_i}{\eta_o\cos\theta_o}, \tag{72}$$

$$P_e=-\frac{\sin\phi_i}{\eta_e\cos\theta_e}, \tag{73}$$

or,

$$\frac{Q_o\zeta_o}{\eta_o}=i, \tag{74}$$

$$\frac{Q_e\zeta_e}{\eta_e}=-i, \tag{75}$$

which ((74) and (75)) is approximately true for vlf skywaves reflected in northern hemisphere. After considerable ado, the Q-L reflection coefficients with the aid of the approximations (69) to (75) and Snell's law,[12]

$$\sin\phi_i=\eta_{o,e}\sin\theta_{o,e}, \tag{76}$$

[12] The validity of Snell's law for complex directions, θ_o and θ_e, is also implied by eqs (36) and (37).

reduce as follows:

$$T_{ee}=\Delta_b^{-1}\{(\eta_o+\eta_e)\ (\cos^2\phi_i-\cos\theta_o\cos\theta_e)+(\eta_o\eta_e-1)\ (\cos\theta_o+\cos\theta_e)\cos\phi_i\}, \tag{77}$$

$$T_{em}=\Delta_b^{-1}\{2\ i\cos\phi_i\ (\eta_o\cos\theta_o-\eta_e\cos\theta_e)\}, \tag{78}$$

$$T_{me}=\Delta_b^{-1}\{2\ i\cos\phi_i\ (\eta_o\cos\theta_e-\eta_e\cos\theta_o)\}, \tag{79}$$

$$T_{mm}=\Delta_b^{-1}\{(\eta_o+\eta_e)\ (\cos^2\phi_i-\cos\theta_o\cos\theta_e)-(\eta_o\eta_e-1)\ (\cos\theta_o+\cos\theta_e)\cos\phi_i\}, \tag{80}$$

where,

$$\Delta_b=(\eta_o+\eta_e)\ (\cos^2\phi_i+\cos\theta_o\cos\theta_e)+(\eta_o\eta_e+1)\ (\cos\theta_o+\cos\theta_e)\cos\phi_i, \tag{81}$$

which formulas agree with those independently derived by Budden [7]. The indexes of refraction of the model ionosphere, η_o and η_e, corresponding to upgoing "ordinary" and "extraordinary" waves can be calculated rigorously as the roots of the quartic in ζ (21). However, as a result of the matrix asymmetry introduced by the Q–L approximation (74) and (75) an interchange of the roots of the quartic (η_o and η_e) results in an ambiguity, $\pm\pi$, in the argument or phase of the abnormal components, arg T_{em} and arg T_{me} (78) and (79). At very low frequencies, the indexes of refraction η_o and η_e are chosen such that arg T_{em} and arg T_{me} are close to π. The values η can also be calculated from the Appleton-Hartree formula (34). Budden [7] and Wait [8] employ additional approximations to this formula,[13]

$$\eta_{o,e}^2=1-i\frac{\omega_r}{\omega}\exp[\pm i\phi_1], \tag{82}$$

in which the plus (+) sign before ϕ_1 refers to η_o and the minus (−) sign refers to η_e and where,

$$\tan\phi_1=\frac{\omega_L}{\nu}, \tag{83}$$

$$\frac{\omega_r}{\omega}=\frac{\omega_{cr}^2}{\omega\sqrt{\nu^2+\omega_L^2}}. \tag{84}$$

6. Computations and Discussion

The results of a calculation [14] of model ionosphere reflection coefficients which can, under certain conditions, represent the reflection of waves from the *D*-region or the *E*-region of the ionosphere are illustrated, figures 3 to 32. These data have been presented as a function of frequency, for various angles of incidence on the ionosphere, ϕ_i. The application of these reflection coefficients to the geometric-optical theory is quite simple,[15] however, the relation between angle of incidence, ϕ_i, distance of observer on the surface of the earth from the source on the surface of the earth, d/j (j=order of hop or time-mode, i.e., j=1, 2, 3 . . .), and the altitude of the boundary of the model ionosphere have been presented for the convenience of the reader, figure 2. The results of a rigorous calculation of the reflection coefficients of the model ionosphere are compared with the results of the Q–L approximation, figures 29 to 32. Thus, for magnetic inclinations near 60°, the Q–L approximation is quite satisfactory at frequencies less than 8 kc for the normal components, T_{ee} and T_{mm}. The most serious discrepancy occurs in the prediction of the abnormal components, $|T_{em}|$ (fig. 29, for example). At higher vlf frequencies ($>$ 8 kc) and at lf, the Q–L theory fails except in certain special cases ($|T_{me}|$. fig. 32, for example).

The effects of magnetic azimuth ϕ_a, figures 19, 20 (compare with figs. 3, 4) magnetic inclination or dip I, figures 21 to 24, magnetic intensity H_m, figures 25 to 28, and distance or angle of incidence ϕ_i, figures 3 to 18, on the reflection coefficients are illustrated. It is of interest to note that an increase in the earth's magnetic intensity H_m increases the abnormal components T_{em}, T_{me}, figures 27, 28. It is also of interest to note that the reciprocity theorem does not

[13] The validity of these approximations has been discussed by Budden [7].
[14] The calculations were performed on the IBM-650 electronic data processing machine.
[15] See for example, H. Bremmer [5] p. 153 et seq.

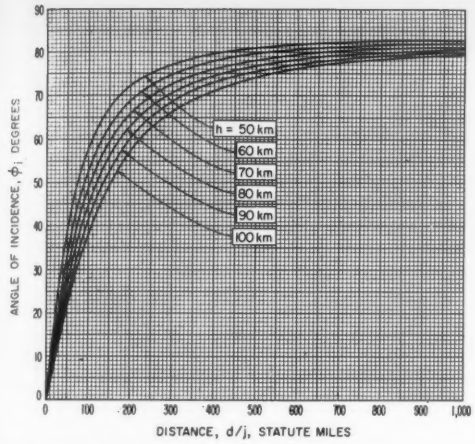


FIGURE 2. *Geometric-optical relation between the angle of incidence, ϕ_i, distance from the source,* d/j, (j=*order of skywave hop,* j=*1, 2, 3 . . .*) *and altitude of boundary of model ionosphere,* h.

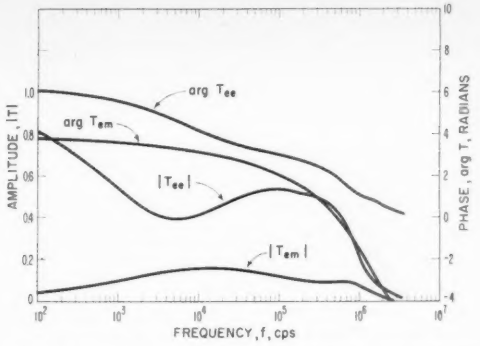


FIGURE 3. *Model ionosphere reflection coefficients;* ϕ_i=*75.08°, sin* ϕ_i=*0.966356,* N=*870,* ν=*4(10⁶),* H_m=*0.5,* ϕ_a=*0,* I=*60° (for example, fig. 2,* d/j=*329 miles,* h=*65 km).*

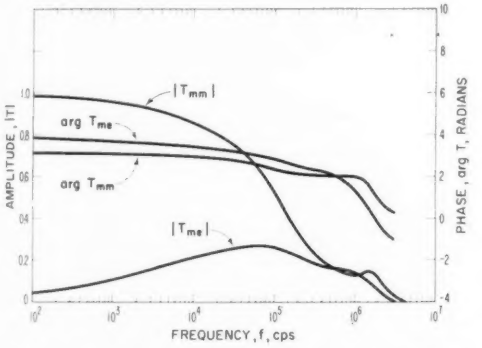


FIGURE 4. *Model ionosphere reflection coefficients;* ϕ_i=*75.08°, sin* ϕ_i=*0.966356,* N=*870,* ν=*4(10⁶),* H_m=*0.5,* ϕ_a=*0,* I=*60° (for example, fig. 2,* d/j=*329 miles,* h=*65 km).*

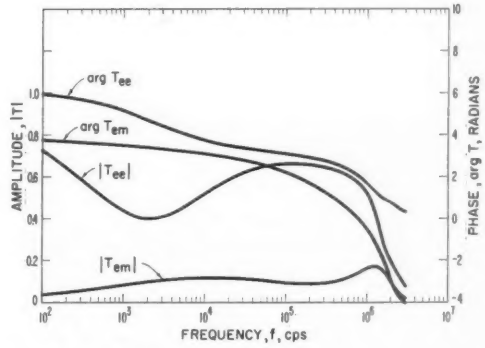


FIGURE 5. *Model ionosphere reflection coefficients;* ϕ_i=*80.39°, sin* ϕ_i=*0.985948,* N=*870,* ν=*4(10⁶),* H_m=*0.5,* ϕ_a=*0,* I=*60° (for example, fig. 2,* d/j=*621 miles,* h=*65 km).*

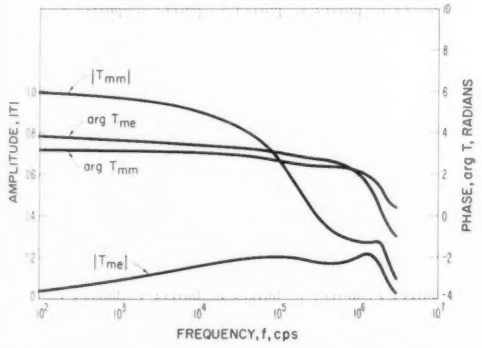


FIGURE 6. *Model ionosphere reflection coefficients;* ϕ_i=*80.39°, sin* ϕ_i=*0.985948,* N=*870,* ν=*4(10⁶),* H_m=*0.5,* ϕ_a=*0,* I=*60° (for example, fig. 2,* d/j=*621 miles,* h=*65 km).*

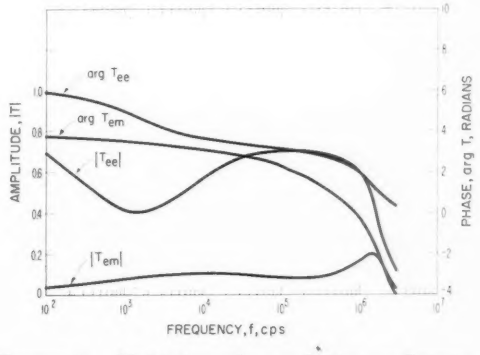


FIGURE 7. *Model ionosphere reflection coefficients;* ϕ_i=*81.79°, sin* ϕ_i=*0.989753,* N=*870,* ν=*4(10⁶),* H_m=*0.5,* ϕ_a=*0,* I=*60° (for example, fig. 2,* d/j=*1000 miles,* h=*65 km).*

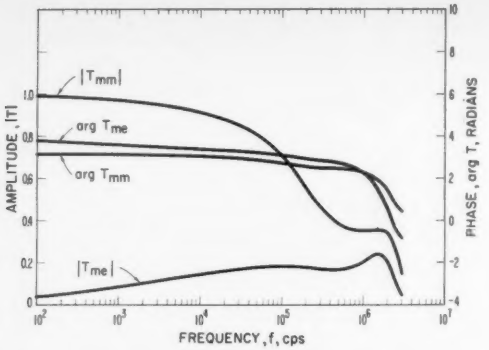


FIGURE 8. *Model ionosphere reflection coefficients;* $\phi_i = 81.79°$, *sin* $\phi_i = 0.989753$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *1000 miles,* h = *65 km).*

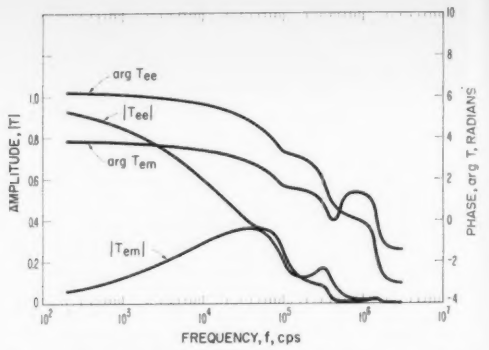


FIGURE 9. *Model ionosphere reflection coefficients;* $\phi_i = 43.25°$, *sin* $\phi_i = 0.685204$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *100 miles,* h = *85 km).*

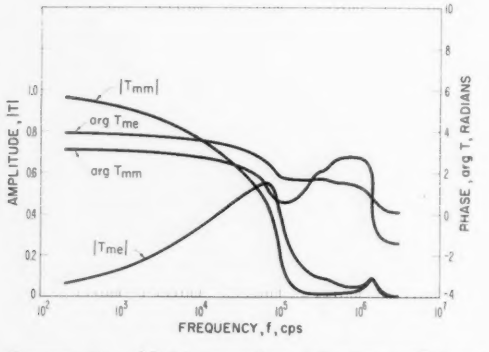


FIGURE 10. *Model ionosphere reflection coefficients;* $\phi_i = 43.25°$, *sin* $\phi_i = 0.685204$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *100 miles,* h = *85 km).*

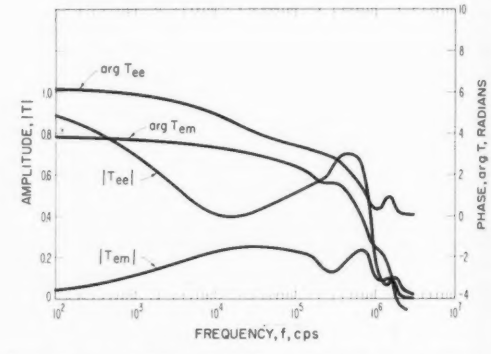


FIGURE 11. *Model ionosphere reflection coefficients;* $\phi_i = 71.13°$, *sin* $\phi_i = 0.946258$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *329 miles,* h = *85 km).*

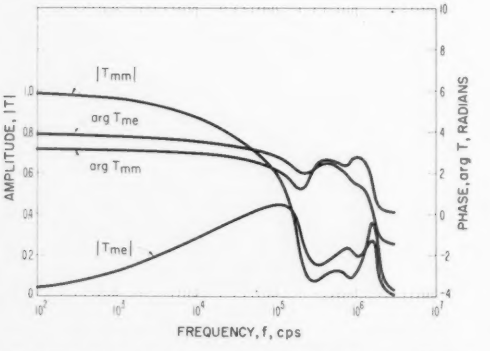


FIGURE 12. *Model ionosphere reflection coefficients;* $\phi_i = 71.13°$, *sin* $\phi_i = 0.946258$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *329 miles,* h = *85 km).*

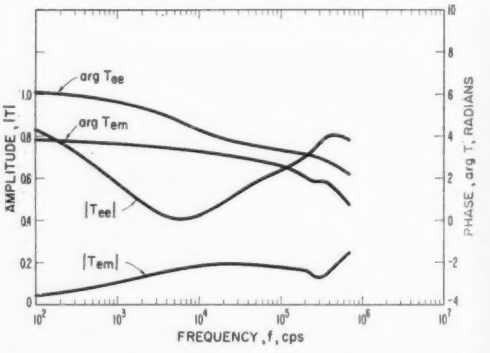


FIGURE 13. *Model ionosphere reflection coefficients;* $\phi_i = 78.17°$, *sin* $\phi_i = 0.978751$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *621 miles,* h = *85 km).*

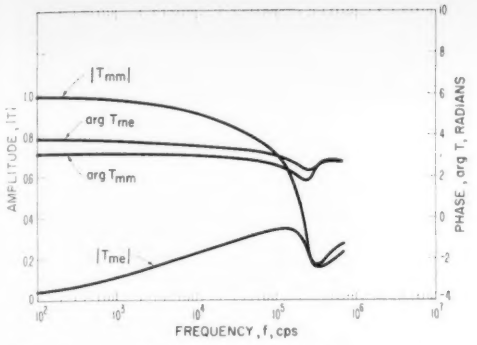


FIGURE 14. *Model ionosphere reflection coefficients;* $\phi_i = 78.17°$, *sin* $\phi_i = 0.978751$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *621 miles,* h = *85 km).*

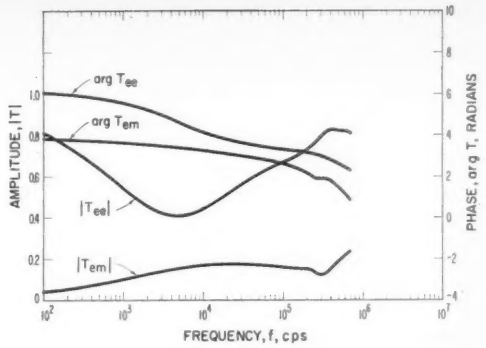


FIGURE 15. *Model ionosphere reflection coefficients;* $\phi_i = 79.64°$, *sin* $\phi_i = 0.983688$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *800 miles,* h = *85 km).*

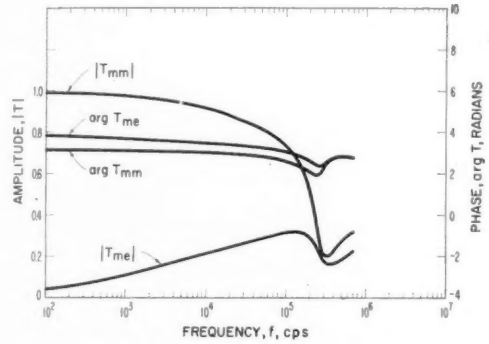


FIGURE 16. *Model ionosphere reflection coefficients;* $\phi_i = 79.64°$, *sin* $\phi_i = 0.983688$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j. = *800 miles,* h = *85 km).*

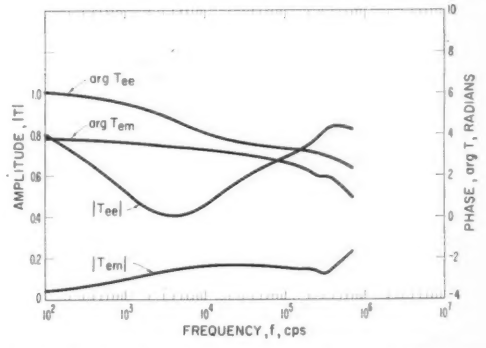


FIGURE 17. *Model ionosphere reflection coefficients;* $\phi_i = 80.40°$, *sin* $\phi_i = 0.985987$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *1000 miles,* h = *85 km).*

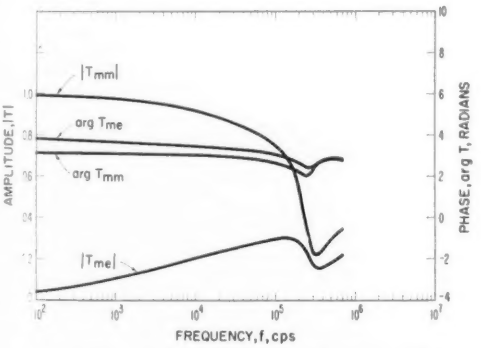


FIGURE 18. *Model ionosphere reflection coefficients;* $\phi_i = 80.40°$, *sin* $\phi_i = 0.985987$, $N = 1200$, $\nu = 10^6$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(for example, fig. 2,* d/j = *1000 miles,* h = *85 km).*

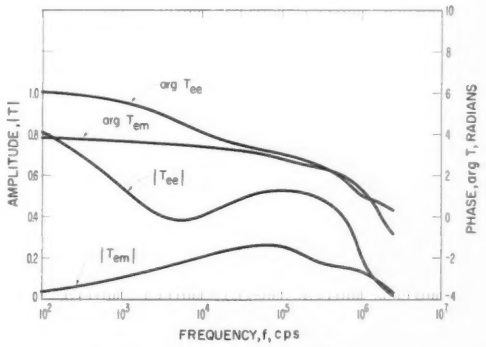


FIGURE 19. *Model ionosphere reflection coefficients, illustrating the effect of magnetic azimuth,* ϕ_a *(compare with fig. 3);* $\phi_i = 75.08°$, *sin* $\phi_i = 0.966356$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.5$, $\phi_a = 180°$, $I = 60°$ *(for example, fig. 2,* d/j = *329 miles,* h = *65 km).*

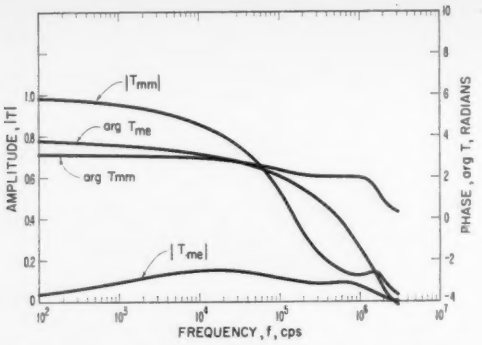


FIGURE 20. *Model ionosphere reflection coefficients, illustrating the effect of magnetic azimuth, ϕ_a (compare with fig. 4); $\phi_i=75.08°$, sin $\phi_i=0.966356$, N=870, $\nu=4(10^6)$, $H_m=0.5$, $\phi_a=180°$, I=60° (for example, fig. 2, d/j=329 miles, h=65 km).*

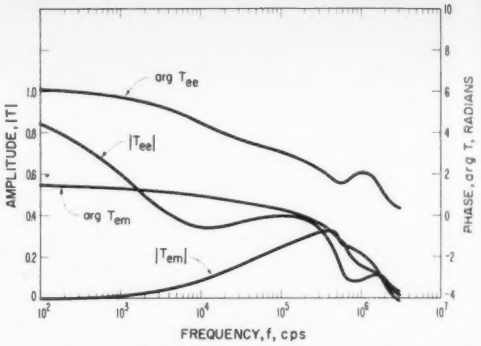


FIGURE 21. *Model ionosphere reflection coefficients, illustrating the effect of magnetic inclination or dip, I (compare with figs. 3, 23); $\phi_i=75.08°$, sin $\phi_i=0.966356$, N=870, $\nu=4(10^6)$, $H_m=0.5$, $\phi_a=0$, I=0.*

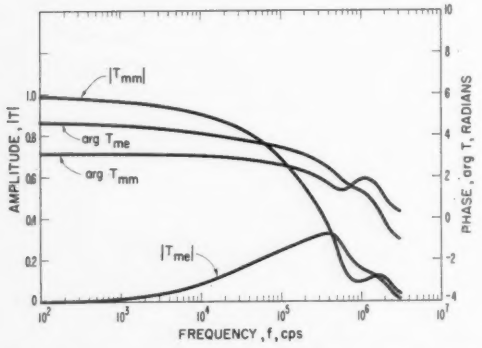


FIGURE 22. *Model ionosphere reflection coefficients, illustrating the effect of magnetic inclination or dip, I (compare with figs. 4, 24); $\phi_i=75.08°$, sin $\phi_i=0.966356$, N=870, $\nu=4(10^6)$, $H_m=0.5$, $\phi_a=0$, I=0.*

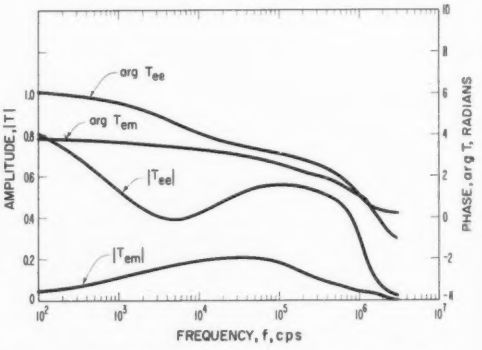


FIGURE 23. *Model ionosphere reflection coefficients, illustrating the effect of magnetic inclination or dip, I (compare with figs. 3, 21); $\phi_i=75.08°$, sin $\phi_i=0.966356$, N=870, $\nu=4(10^6)$, $H_m=0.5$, $\phi_a=0$, I=90°.*

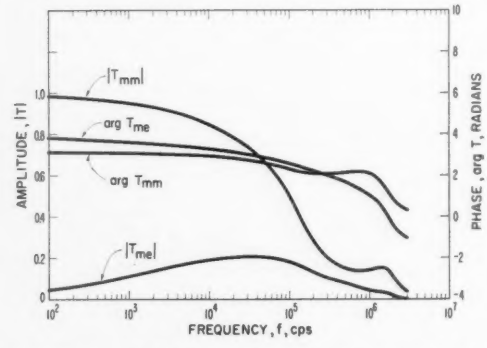


FIGURE 24. *Model ionosphere reflection coefficients, illustrating the effect of magnetic inclination or dip, I (compare with figs. 4, 22); $\phi_i=75.08°$, sin $\phi_i=0.966356$, N=870, $\nu=4(10^6)$, $H_m=0.5$, $\phi_a=0$, I=90°.*

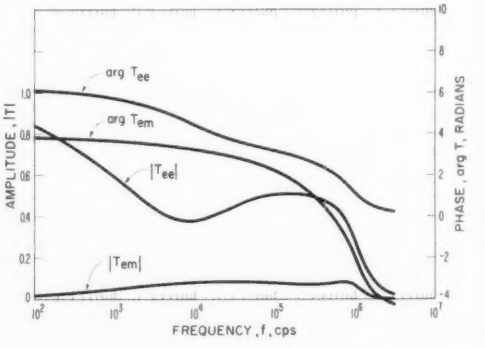


FIGURE 25. *Model ionosphere reflection coefficients, illustrating the effect of the intensity, H_m, of the earth's magnetic field vector, $\overline{H}_m$ (compare with figs. 3, 27); $\phi_i=75.08°$, sin $\phi_i=0.966356$, N=870, $\nu=4(10^6)$, $H_m=0.2$, $\phi_a=0$, I=60°.*

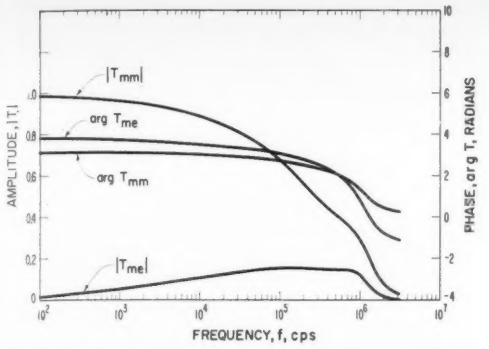


FIGURE 26. *Model ionosphere reflection coefficients, illustrating the effect of the intensity,* H_m, *of the earth's magnetic field vector,* $\overline{H}_m$ *(compare with figs. 4, 28);* $\phi_i = 75.08°$, *sin* $\phi_i = 0.966356$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.2$, $\phi_a = 0$, $I = 60°$.

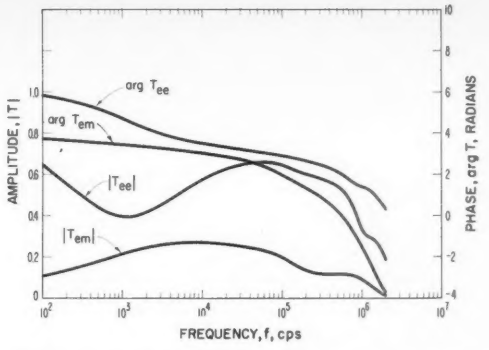


FIGURE 27. *Model ionosphere reflection coefficients, illustrating the effect of the intensity,* H_m, *of the earth's magnetic field vector,* $\overline{H}_m$ *(compare with figs. 3, 24);* $\phi_i = 75.08°$, *sin* $\phi_i = 0.966356$, $N = 870$, $\nu = 4(10^6)$, $H_m = 1$, $\phi_a = 0$, $I = 60°$.

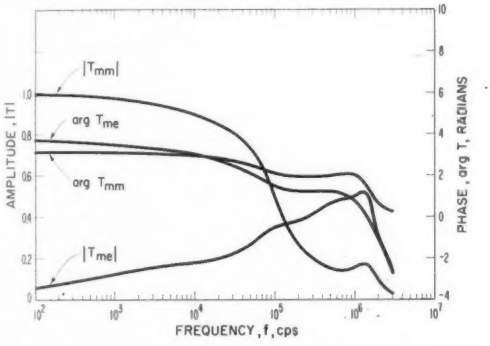


FIGURE 28. *Model ionosphere reflection coefficients, illustrating the effect of the intensity,* H_m, *of the earth's magnetic field vector,* $\overline{H}_m$ *(compare with figs. 4, 25)* $\phi_i = 75.08°$, *sin* $\phi_i = 0.966356$, $N = 870$, $\nu = 4(10^6)$, $H_m = 1$, $\phi_a = 0$, $I = 60°$.

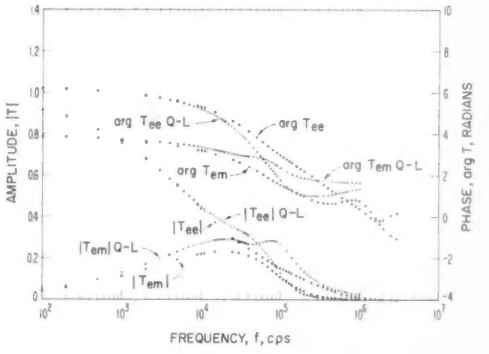


FIGURE 29. *Comparison of reflection coefficients calculated by the rigorous method for the model ionosphere with the Q–L approximation;* $\phi_i = 50.84°$, *sin* $\phi_i = 0.775353$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(*ϕ_a *and* I *do not apply to Q–L method; also note: fig. 2,* d/j *= 100 miles,* h *= 65 km).*

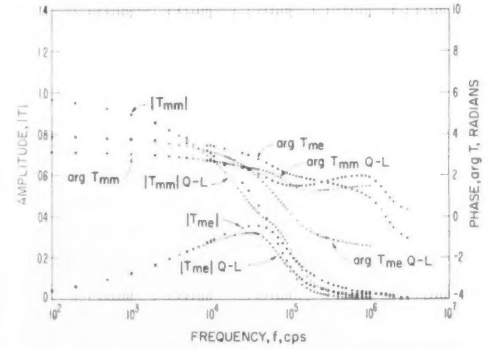


FIGURE 30. *Comparison of reflection coefficients calculated by the rigorous method for the model ionosphere with the Q–L approximation;* $\phi_i = 50.84°$, *sin* $\phi_i = 0.775353$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(*ϕ_a *and* I *do not apply to the Q–L method; also note: fig. 2,* d/j *= 100 miles,* h *= 65 km).*

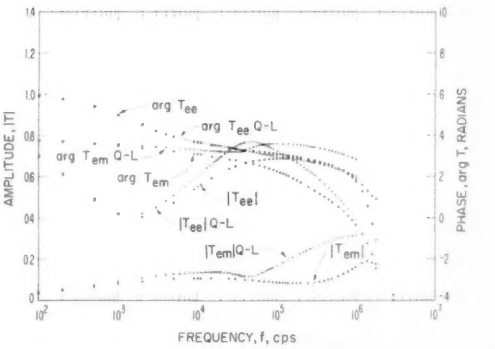


FIGURE 31. *Comparison of reflection coefficients calculated by the rigorous method for the model ionosphere with the Q–L approximation;* $\phi_i = 81.37°$, *sin* $\phi_i = 0.988681$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ *(*ϕ_a *and* I *do not apply to the Q–L method; also note: fig. 2,* d/j *= 800 miles,* h *= 65 km).*

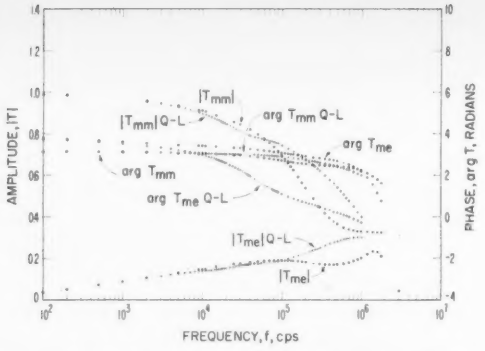


FIGURE 32. *Comparison of reflection coefficients calculated by the rigorous method for the model ionosphere with the Q–L approximation;* $\phi_i = 81.37°$, *sin* $\phi_i = 0.988681$, $N = 870$, $\nu = 4(10^6)$, $H_m = 0.5$, $\phi_a = 0$, $I = 60°$ (ϕ_a *and* I *do not apply to the Q–L method; also note: fig. 2,* d/j = *800 miles,* h = *65 km*).

hold, i.e., south-north propagation, $\phi_a = 0$, figures 3, 4, is not precisely equivalent to north-south propagation, $\phi_a = 180°$, figures 19, 20. It is important to note that the magnetic dip I, figures 21 to 24, and the magnetic azimuth ϕ_a, figures 19 to 20, i.e., the direction of the earth's magnetic field vector $\overline{H}_m$ is a very significant consideration, not only at lf but also at vlf.

Since the actual ionosphere electron [16] density versus altitude profile is quite variable, considerable caution is necessary in the application of this model to the interpretation of experimental data. It can, in general, be stated that the model is valid when the boundary of the ionosphere is sufficiently sharp (relative to a wavelength) at the frequency of the incident wave under consideration. The precision with which this model represents the actual ionosphere can be considered a first approximation at low frequencies, the degree of approximation decreasing with increasing frequency.

7. Conclusions

The reflection coefficients for the sharply bounded model ionosphere can be readily evaluated with the aid of electronic data processing techniques for the solution of the quartic equation and the tensor matrix, thus eliminating the necessity for the Q–L approximation. The Q–L theory is a crude mathematical approximation taking account of the intensity but neglecting completely the direction (magnetic azimuth and dip) of the earth's magnetic field vector and if used must be applied with considerable caution not only at lf but also at vlf frequencies. The complete and rigorous magneto-ionic solution for the assumed model ionosphere not only eliminates the necessity for such an approximation but also illustrates the application of modern numerical analysis techniques which can in many cases obviate such mathematical approximations, leaving only the physical approximations implied by the assumed model.

The model ionosphere presented in this paper represents a first approximation of the electron density-altitude profile, the degree of approximation decreasing with increasing frequency and the validity of the model dependent upon the sharpness of the actual ionosphere boundary relative to a wavelength. The results of this paper suggest a refinement of the theory which would take into account a "less sharply bounded" model ionosphere.

[16] Values of electron density, N, and collision frequency, ν, employed in this paper were estimated from recent data, see for example, Waynick [13].

8. References

[1] E. V. Appleton, Wireless studies of the ionosphere, J. Inst. Elec. Engrs. (London) **71,** 642 (1932).

[2] D. R. Hartree, The propagation of electromagnetic waves in a stratified medium, Proc. Cambridge Phil. Soc. **25,** 97 (1929)

[3] D. R. Hartree, The propagation of electromagnetic waves in a refracting medium in a magnetic field, Proc. Cambridge Phil. Soc. **27,** 143 (1931).

[4] H. G. Booker, Some general properties of the formulae of the magneto-ionic theory, Proc. Roy. Soc. (London) **A147,** 352 (Nov. 1934).

[5] H. Bremmer, Terrestrial radio waves—theory of propagation (Elsevier Publishing Co., New York, N.Y., 1949).

[6] I. W. Yabroff, Reflection at a sharply bounded ionosphere, Proc. IRE **45,** No. 6, 750 (June 1957). See also I. W. Yabroff, Reflection and transmission at a sharply bounded ionosphere, Tech. Rept. No. 1, Radio Prop. Lab., Stanford Univ., Stanford, Calif. (July 29, 1957).

[7] K. G. Budden, The reflection of very low frequency radio waves at the surface of a sharply bounded ionosphere with superimposed magnetic field, Phil. Mag. 2, **42,** 833 (Aug. 1951).

[8] J. R. Wait and L. B. Perry, Calculation of ionosphere reflection coefficients at very low frequencies, J. Geophys. Research **1,** No. 1, 62 (March 1957).

[9] D. E. Muller, A method for solving algebraic equations using an automatic computer, Math. Tables and Other Aids to Computation **10,** No. 53–56, 208 (1956).

[10] F. B. Hildebrand, Introduction to numerical analysis, p. 60 (McGraw-Hill Book Co., Inc., New York, N.Y., 1956).

[11] J. A. Stratton, Electromagnetic theory (McGraw-Hill Book Co., Inc., New York, N.Y., 1941).

[12] P. D. Crout, Determinants and solving systems of linear equations with real or complex coefficients, Proc. AIEE, **60,** 1235 (1941).

[13] A. H. Waynick, The present state of knowledge concerning the lower ionosphere, Proc. IRE **45,** No. 6 (June 1957).

Boulder, Colo.

(Paper 64D3–58)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

Focusing, Defocusing, and Refraction in a Circularly Stratified Atmosphere[1]

K. Toman[2]

(October 29, 1959)

Focusing, defocusing, astronomical refraction and path length of rays as a function of the departure angle Δ of the ray at the source is described for cases with the source outside, inside, or on the boundary of a circular stratification. Relative to zero elevation angle symmetrical and centrosymmetrical distributions are found.

1. Introduction

In order to determine focusing, refraction, and path-length properties of rays in a circularly stratified atmosphere for all departure angles at a source, the earth is temporarily removed. If the earth is restored the use of the obtained distributions is unaffected for positive departure angles, but the use of negative departure angles is limited by the elevation of the source. Applied to a spherical lense system, however, the distributions apply for all departure angles.

2. Focusing and Defocusing

Place an electromagnetic emitter (P at $r=C\neq 0$) in a medium of constant refractive index (μ_1) which is circularly bounded. Designate $\mu_1(\leq r_1)=1$, $\mu_2(r_1<r<r_2)<1$, and $\mu_3(\geq r_2)=1$. Define zero elevation angle as the direction s which is normal to the diameter through P. Let the smaller quadrant contain positive elevation angles. Assume an opening angle $\delta\Delta$ and let the beam rotate on P. The behavior of the opening angle of the refracted beam shows that for positive elevation angles the beam is defocused. For negative elevation angles the beam is focused. If $\mu_2(r_1<r<r_2)>1$ these conditions are reversed. If the beam is centered in a direction $\Delta=0$ there is neither focusing nor defocusing. Geometrically, this is true for any beam width. For rays the beam width is limited by the known restrictions of ray theory.

These results apply to a circularly stratified ionosphere without magnetic field for which $\mu_2(r)<1$ if the influence of the collisional frequency is neglected. Taking into account collisional frequency, the refractive index can become greater than unity and the net effect is either focusing or defocusing. If, however, defocusing is obtained for a particular value of Δ defocusing strictly prevails over the full quadrant to which Δ belongs. Here, rays are assumed to originate at P. Should the source be at the location of a circularly orbiting satellite monitored from P the reciprocity theorem can be invoked.

To illustrate these effects, the differential refraction angle $\delta\epsilon$ of a beam with a beam width $\delta\Delta=5°$ passing through a medium $\mu_1=\mu_3=1$, $\mu_2=0.92$ is shown in figure 1. The $\delta\epsilon(\Delta)$ distribution was obtained using convenient values for the radial distance of P at $C=90$ with $r_1=100$ and $r_2=110$. Before entering the refracting medium rays are linear. At $\Delta=90°$ (zenith angle $z=0$) there remains a net amount of defocusing which is zero for zero beam width. For smaller values of Δ defocusing is enhanced until a maximum is reached. For $\Delta=0$ there is neither focusing nor defocusing. The $\delta\epsilon(\Delta)$ distribution is centrosymmetrical relative to $\Delta=0$. Similar distributions are obtained if μ_2 varies with

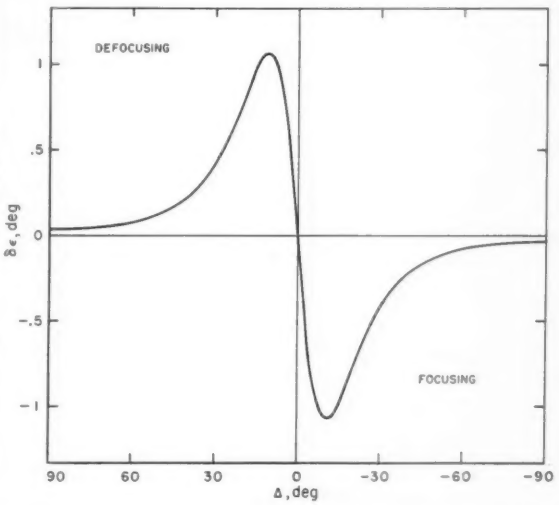


FIGURE 1. *Focusing and defocusing for a circularly stratified medium.*

[1] Presented, in parts, at the spring URSI meeting, May 5 to 7, 1959, Washington, D. C.

[2] Geophysics Research Directorate, Ionospheric Physics Laboratory, Air Force Cambridge Research Center, Bedford, Mass. (invited paper).

height in an arbitrary manner. The $\delta\epsilon$ (Δ) distribution remains centrosymmetrical if P lies at the boundary of stratification provided at this boundary μ is single valued. If P lies inside the stratification between r_1 and r_2, the $\delta\epsilon$ (Δ) distribution no longer retains its centrosymmetry because for negative departure angles at P the amount of differential refraction depends on $\mu(r)$ for $r<C$.

3. Refraction

For the same parameters of figure 1 the astronomical refraction $\delta\eta$ was obtained as shown in figure 2. Maximum deviation of the ray from its initial direction at P occurs at $\Delta=0$. An interesting case is $\delta\eta=0$ for $\Delta\neq 90°$. Then $\delta\eta$ is zero for all values of Δ. Consequently, $\delta\epsilon$ is zero and zero-focusing prevails for all departure angles at P.

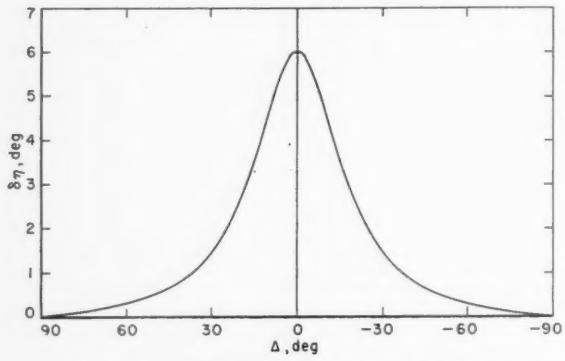


FIGURE 2. *Astronomical refraction $\delta\eta$ for a circularly stratified medium.*

The symmetry of $\delta\eta$ is valid for a circularly stratified atmosphere of any $\mu(r)$ distribution provided the lowest stratification level is at $r_1\geq C$ and μ is single valued at $r_1=C$. For $r_1=C$ the $\delta\eta$ (Δ) distribution will no longer have zero gradient at $\Delta=0$ since the behavior of $\delta\eta$ depends on the gradient of μ at the departure point P. If μ at $r_1=C$ is double valued the $\delta\eta$ (Δ) distribution ceases to be symmetrical. For reasons indicated above the $\delta\eta$ (Δ) distribution also ceases to be symmetrical if P lies inside the stratification.

To make use of the symmetry a radio source must be located there where μ does not change with height. These conditions may best be fulfilled in the space between the troposphere and the ionosphere. At such a location one obtains extremes of $\delta\epsilon$ (Δ) and $\delta\eta$ (Δ) at $\Delta=0$. These distributions are limited by a negative elevation angle for which the ray grazes the "top" of the troposphere.

Snell's law for spherical refracting surfaces derives the orientation of the ray from the boundary values of the $\mu(r)$ distribution. For a circularly stratified atmosphere it is thus possible to determine the $\delta\eta(\Delta)$ distribution for any $\mu(r)$ distribution in a simple manner. First, one chooses a departure angle $\Delta\neq 90°$ and traces the ray path from r_1 to r_2. Between r_1 and r_2 one selects μ=constant such that the linear ray intersects the actual ray path at r_1 and r_2. Outside r_2, μ is chosen so that the angle of intersection of the ray with the radius vector is the same. For this model the complete $\delta\eta(\Delta)$ distribution is easily computed.

4. Ray-Path Length

Symmetry properties with respect to zero elevation angle Δ can also be stated in a general form for the length of ray paths which originate at P and which traverse a circularly stratified medium. If P lies below the boundary of stratification the path length between r_1 and r_2 of the refracted ray is a maximum for zero departure angle of the ray at P.[3] The path-length distribution is symmetrical relative to $\Delta=0$ at P. If P lies on the boundary of stratification ($C=r_1$) this symmetry is preserved provided μ at C is single valued. If P lies inside the stratification the symmetry of the total path length vanishes but the line integrals between fixed levels for the geometrical path length

$$\int_{r\geq C}^{C<r\leq\infty} ds$$

and for the optical path length

$$\int_{r\geq C}^{C<r\leq\infty} \mu\, ds$$

of rays emerging from a source have both a maximum for zero departure angle of the ray at P. Their values are symmetrical relative to $\Delta=0$.

Applied to a circularly stratified ionosphere for $C\leq r_1$, these results illustrate that the maximum of a nondeviative type of absorption is strictly obtained for zero departure angle of the ray at the source.

5. Conclusion

If a source of rays (P at $r=C$) lies below or on the boundary of a circular stratification of the refractive index, the astronomical refraction and the total ray path length are both symmetrical relative to $\Delta=0$ with a maximum at $\Delta=0$; the differential refraction is centrosymmetrical. If P lies inside the stratification, these symmetries vanish although the line integrals of the geometrical and the optical path length between levels of $r>C$ are symmetrical and obtain a maximum at $\Delta=0$.

[3] K. Toman, New geometrical properties and their usefulness for ionospheric radio propagation, Proc. IRE **47**, 1381 (1959).

BOULDER, COLO. (Paper 64D3-59)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

Response of a Loaded Electric Dipole in an Imperfectly Conducting Cylinder of Finite Length [1]

Charles W. Harrison, Jr.,[2] and Ronold W. P. King [3]

(November 16, 1959)

Analytical relationships are developed which permit calculation of the power in the load impedance of an electric probe, symmetrically located within an imperfectly conducting cylinder of small radius compared to the wavelength, in terms of the electric field incident upon the cylinder.

1. Introduction

The problem is to determine the power in the load impedance at the center of a short dipole receiving antenna that is located within an imperfectly conducting shield of cylindrical shape and finite length in terms of a known electric field outside the completely closed shield. It is assumed that the radius of the shield and length of the probe are small fractions of the wavelength. The solution consists of the successive determination of the following:

(a) The ratio of the electric field inside to the field on the outside surface of the shield.

(b) The ratio of the electric field on the outside surface of the hollow cylinder to the incident electric field.

(c) The effective length and the driving-point impedance of the center-loaded dipole inside the shield.

(d) The power in the load impedance of the dipole in terms of the electric field within the cylindrical shield.

2. Ratio of the Electric Field Inside an Imperfectly Conducting Shield to the Field on the Outside Surface [1] [4]

Figure 1 shows a tube with hemispherical end caps. Its inner radius is a_1, its outer radius a_2, and its axial length $2l$. The metal annulus is region 1, the spaces outside and inside the tube are regions 2 and 3, respectively. Let the axis of the cylindrical shield coincide with the z axis of a system of cylindrical coordinates r, θ, and z. If the length of the tube is at least 10 times its diameter and $a_2 << \lambda$, it is a satisfactory approximation to assume the transverse distribution of the axial electric field at any cross section of the cylinder not too near its ends to be independent of its axial distribution. The field outside the cylinder is assumed to be directed parallel to its axis, i.e., in the z-direction.

The governing wave equation for the electric field $\underset{\sim}{E}=\hat{z}E_z$ in a homogeneous region with conductivity $\sigma_1 >> \omega\epsilon_0$, dielectric constant $\epsilon_1=\epsilon_0$, and permeability μ_1 is

$$V^2\underset{\sim}{E}-j\omega\sigma_1\mu_1\underset{\sim}{E}=0. \tag{1}$$

A time dependence of the form $\exp(j\omega t)$ is assumed where $\omega=2\pi f$ and f is the frequency.

Owing to rotational symmetry and the assumed independence of the transverse and axial distributions, the equation for the transverse distribution of E_z is

$$\frac{\partial^2 E_z(r)}{\partial r^2}+\frac{1}{r}\frac{\partial E_z(r)}{\partial r}+k_m^2 E_z(r)=0, \tag{2}$$

which has the general solution

$$E_z(r)=A_1 J_0(k_m r)+B_1 N_0(k_m r) \qquad a_1 \le r \le a_2, \tag{3}$$

where

$$k_m=(1-j)\sqrt{\frac{\omega\sigma_1\mu_1}{2}}, \tag{4}$$

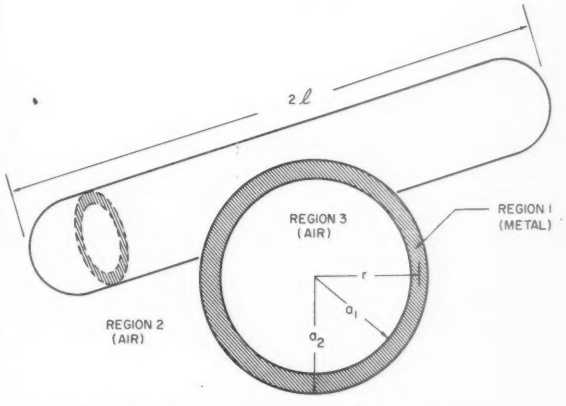


Figure 1. Cross section of hollow conductor.

[1] This paper originally appeared as Sandia Corporation Technical Memorandum No. 457-58(14) dated December 5, 1958.

[2] Member of the Scientific Staff of the Sandia Corporation, Sandia Base, Albuquerque, New Mexico.

[3] Consultant to the Sandia Corporation and Gordon McKay Professor of Applied Physics, Harvard University.

[4] Figures in brackets indicate the literature references at the end of this paper.

and $J_0(x)$ and $N_0(x)$ are Bessel functions of the first and second kind. In region 3 the free space wave equation,

$$\nabla^2 \underset{\sim}{E} + \omega^2 \mu_0 \epsilon_0 \underset{\sim}{E} = 0, \tag{5}$$

applies; $\mu_0 = 4\pi \times 10^{-7} h/m$; $\epsilon_0 = 8.85 \times 10^{-12} f/m$. The solution of (5) that remains finite at $r=0$ is

$$E_z(r) = A_3 J_0(k_s r) \qquad 0 \leq r \leq a_1, \tag{6}$$

where

$$k_s = \omega \sqrt{\mu_0 \epsilon_0}. \tag{7}$$

Now, if $(k_s a_1)^2 << 1$, $J_0(k_s r) \approx 1$, and $E_z(r) \approx A_3 = \text{const}$ $(0 \leq r \leq a_1)$. It follows that within the cylindrical shell

$$\frac{\partial E_z(r)}{\partial r} = 0. \tag{8}$$

The Maxwell equation

$$\text{curl}_\theta \underset{\sim}{E} = -\frac{\partial E_z}{\partial r} = -j\omega B_\theta \tag{9}$$

when combined with (8) leads to the conclusion that

$$B_\theta(r) = 0 \qquad (0 \leq r \leq a_1). \tag{10}$$

Since the B-vector satisfies the boundary condition

$$\mu_0 B_{\theta 3}(a_1) = \mu_1 B_{\theta 1}(a_1) \tag{11}$$

at $r=a_1$, it follows with (8), (9), and (11) that

$$j\omega B_{\theta 1}(a_1) = \left[\frac{\partial E_{z1}(r)}{\partial r}\right] = 0 \qquad r = a_1. \tag{12}$$

With (12), the constant B_1 may be evaluated in terms of A_1 from (3). The result is

$$B_1 = -A_1 \frac{J_1(k_m a_1)}{N_1(k_m a_1)}. \tag{13}$$

The substitution of (13) in (3) and its evaluation successively at r and $r=a_2$ leads directly to

$$\frac{E_z(r)}{E_z(a_2)} = \frac{J_0(k_m r) N_1(k_m a_1) - N_0(k_m r) J_1(k_m a_1)}{J_0(k_m a_2) N_1(k_m a_1) - N_0(k_m a_2) J_1(k_m a_1)}$$

$$a_1 \leq r \leq a_2. \tag{14}$$

Finally, when $r=a_1$

$$\frac{E_z(a_1)}{E_z(a_2)} = \frac{J_0(k_m a_1) N_1(k_m a_1) - N_0(k_m a_1) J_1(k_m a_1)}{J_0(k_m a_2) N_1(k_m a_1) - N_0(k_m a_2) J_1(k_m a_1)}, \tag{15}$$

which is the desired complex ratio of the field inside the imperfectly conducting hollow cylinder to the field on its outside surface.

In the important special case of large arguments defined by [5]

$$a_1 \sqrt{\omega \mu_1 \sigma_1} \geq 10, \tag{16}$$

it is possible to use the well known asymptotic forms of the Bessel functions to simplify (15). The result is

$$\frac{E_z(a_1)}{E_z(a_2)} = \sqrt{\frac{a_2}{a_1}} \left\{ \frac{1}{\cos\left[k_m(a_2 - a_1)\right]} \right\}. \tag{17}$$

With (4) and standard formulas for the cosine with a complex argument, (17) becomes

$$\frac{E_z(a_1)}{E_z(a_2)} = \sqrt{\frac{a_2}{a_1}} \left\{ \frac{\cos\gamma \cosh\gamma - j \sin\gamma \sinh\gamma}{\cos^2\gamma + \sinh^2\gamma} \right\}, \tag{18}$$

and its magnitude is

$$\frac{|E_z(a_1)|}{|E_z(a_2)|} = \sqrt{\frac{a_2}{a_1}} \left\{ \frac{1}{\sqrt{\cos^2\gamma + \sinh^2\gamma}} \right\}, \tag{19}$$

where

$$\gamma = (a_2 - a_1) \sqrt{\frac{\omega \sigma_1 \mu_1}{2}}. \tag{20}$$

Note that (18) and (19) are valid only when (16) is satisfied.

3. Relation Between the Field Incident Upon an Imperfectly Conducting Hollow Cylinder and the Field on its Outside Surface

The volume density of current $i_z(r)$ in the conducting tube $(a_1 \leq r \leq a_2)$ is related to the electric field $E_z(r)$ by the constitutive relation

$$i_z(r) = \sigma_1 E_z(r). \tag{21}$$

It follows that

$$\frac{i_z(r)}{i_z(a_2)} = \frac{E_z(r)}{E_z(a_2)}. \tag{22}$$

The total axial current $I(0)$ at the center $(z=0)$ of the cylinder is given by the formula

$$I(0) = \int_{a_1}^{a_2} i_z(r) 2\pi r dr, \tag{23}$$

where $i_z(r)$ is the transverse distribution of the current density across the center of the cylinder. The substitution of (22) with (14) in (23) and the subsequent integration gives the following expression for

[5] At 15 kc an aluminum cylinder must be at least 0.447 cm in radius; an iron shield could be smaller.

the current density near the outer surface in terms of the total current:

$$i_z(a_2)=\frac{I(0)}{\pi a_2^2}\left(\frac{k_m a_2}{2}\right)$$

$$\times\left\{\frac{J_0(k_m a_2)N_1(k_m a_1)-N_0(k_m a_2)J_1(k_m a_1)}{J_1(k_m a_2)N_1(k_m a_1)-N_1(k_m a_2)J_1(k_m a_1)}\right\}. \quad (24)$$

The corresponding asymptotic formula for $a_1\sqrt{\omega\mu_1\sigma_1}\geq 10$ is

$$i_z(a_2)=\frac{I(0)}{\pi a_2^2}\left(\frac{k_m a_2}{2}\right)\cot\left[k_m(a_2-a_1)\right], \quad (25)$$

so that with (21) at $r=a_2$

$$E_z(a_2)=\frac{I(0)}{\pi\sigma_1 a_2^2}\left(\frac{k_m a_2}{2}\right)\cot\left[k_m(a_2-a_1)\right]. \quad (26)$$

This expression relates the electric field on the outside surface of the hollow cylinder halfway between its ends to the total current $I(0)$ traversing this cross section of the tube.

The axial distribution $I(z)$ of the current in the cylinder, in particular, its value $I(0)$ at the center, is known from antenna theory since the cylindrical shield is, in effect, a receiving antenna without load (refer to fig. 2). The current at its center is given by [2]

$$I(0)=\frac{V_{oc}}{Z_{in}}=\frac{2l_e E_z^i}{Z_{in}}, \quad (27)$$

where V_{oc} is the open-circuit voltage given by the product of the effective length $2l_e$ of the cylinder multiplied by the incident tangentially directed electric field E_z^i. Z_{in} is the driving-point impedance of a symmetrical center-driven antenna with dimensions identical to those of the shield. The effective length $2l_e$ of an electrically short antenna ($k_s l\leq 0.5$) is given approximately by [3]

$$2l_e\approx l, \quad (28)$$

where l is the half-length of the shield and $k_s=2\pi/\lambda$.

The impedance of a perfectly conducting short antenna [3] for which $k_s l\leq 0.5$ is given by the formula [4]

$$Z_0^r\approx -j\frac{\eta}{2\pi k_s l}$$

$$\times\left\{(\Omega-2-2\ln 2)\left(1-j\frac{(k_s l)^3}{3(\Omega-2-2\ln 2)}\right)\right\}\text{ohms}, \quad (29)$$

where

$$\eta=\sqrt{\frac{\mu_0}{\epsilon_0}}=120\pi, \quad (30)$$

and

$$\Omega=2\ln\left(\frac{2l}{a_2}\right). \quad (31)$$

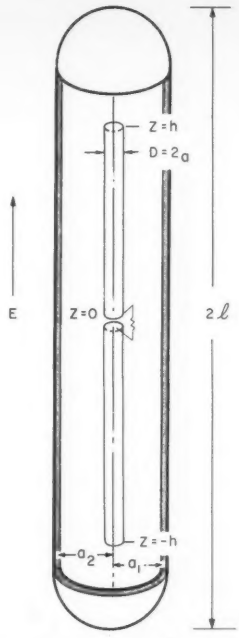


Figure 2. Electric probe symmetrically oriented along the axis of an imperfectly conducting cylinder.

The input impedance of the imperfectly conducting short dipole is approximately.

$$Z_{in}=Z_0^r+R_0^l, \quad (32)$$

where R_0^l is the ohmic loss resistance referred to the input terminals.

R_0^l may be evaluated by noting that the current distribution along an electrically short antenna is approximately triangular in shape. Hence, the averagec urrent along the structure is $I_{av}^2=\frac{1}{3}I^2(0)$. The ohmic resistance of the tube of length $2l$ has been given by King [5]. It is

$$R^t=\frac{lk_m}{\pi a_2\sigma_1}\left(\sqrt{\frac{\cosh 2\gamma+\cos 2\gamma}{\cosh 2\gamma-\cos 2\gamma}}\right)\cos\left(\psi_r-\psi_c+\frac{\pi}{4}\right), \quad (33)$$

where

$$\tan\psi_r=-\tanh\gamma\tan\gamma \quad (34)$$

and

$$\tan\psi_c=\frac{\tanh\gamma}{\tan\gamma} \quad (35)$$

provided $a_1\sqrt{\omega\mu_1\sigma_1}\geq 10$. γ is defined by (20). The power lost in heating the antenna is

$$P_L=I^2(0)R_0^l=\left(\frac{I^2(0)}{3}\right)R^t. \quad (36)$$

Hence,

$$R_0^l = \frac{R^t}{3} \tag{37}$$

Accordingly,

$$Z_{in} \approx -j\frac{\eta}{2\pi k_s l} \times \left\{ (\Omega - 2 - 2\ln 2)\left(1 - j\frac{(k_s l)^3}{3(\Omega - 2 - 2\ln 2)}\right)\right\} + \frac{R^t}{3}. \tag{38}$$

The substitution of (27) in (26) gives

$$E_z(a_2) = \frac{2l_e E_z^i}{\pi Z_{in}\sigma_1 a_2^2}\left(\frac{k_m a_2}{2}\right)\cot\left[k_m(a_2 - a_1)\right]. \tag{39}$$

Thus, if the incident field E_z^i in the vicinity of the cylinder is known, (39) may be used to find the field $E_z(a_2)$ on the outside surface. The field within the cylinder may then be calculated from (18).

4. Effective Height and Driving-Point Impedance of a Short Electric Dipole in an Imperfectly Conducting Cylinder

To determine the behavior of a center-loaded electric dipole of half-length h and radius a symmetrically located along the axis of an imperfectly conducting cylinder requires not only a knowledge of the electric field inside the cylinder in terms of the incident electric field but also the effective length and driving-point impedance of the probe.

The effective length of a probe encased within an imperfectly conducting cylinder is the same as its effective length when isolated. The effective length is defined in terms of the field maintained along the probe by all currents other than those in the probe itself. This is precisely the field $E_z(a_2)$ maintained in the cylinder by the distant signal source and the current in the imperfectly conducting shield. (It is to be remembered that the field in the interior of the tube is uniform.) The only approximation involved is the assumption that the current in the probe does not react on currents in the shield to change them significantly. The required effective length $2h_e$ may be calculated from (28) (or the formula in footnote 3) provided h is written for l throughout (h is the half-length of the electric probe).

The driving-point impedance of the electric probe encased in the imperfectly conducting cylinder lies between the impedance of an identical unshielded probe and the impedance of two imperfectly conducting sections of coaxial transmission-line in series. These sections are open-circuited, and have a length equal to the half-length h of the probe. Capacitive end effects may be determined but are probably small enough to neglect. In terms of the electromagnetic field that would be set up by the probe antenna if it were used for transmission, three situations must be considered:

(a) When the probe is isolated, i.e., is outside the metal cylinder, the distant field is due entirely to the currents in the antenna.

(b) When the probe is in a perfectly conducting cylinder, the field maintained at outside points by the currents in the antenna is exactly cancelled by the field set up by the currents in the shield.

(c) When the probe is in an imperfectly conducting cylinder, currents are induced in the shield which maintain a field that partially cancels the field of the probe, but this cancellation is not complete.

Thus the current in the probe can be separated into two parts. One part, I_L, is equal and opposite to the current in the shield; it is a transmission-line current. The other part, I_A, is the true antenna current. The total current I_t in the unloaded probe is

$$I_t = I_A + I_L. \tag{40}$$

The voltage driving the probe is

$$V_o = I_t Z_0 = I_A Z_A + I_L Z_L = I_t \left\{ Z_L - \frac{I_A}{I_t}(Z_L - Z_A)\right\}, \tag{41}$$

where Z_0 is the impedance of the probe when encased in the imperfectly conducting cylinder; Z_L is the impedance of two perfectly conducting open-circuited sections of coaxial line that are of equal length and connected in series. Z_L may be computed from the formula

$$Z_L = -j\,276\log_{10}\left(\frac{a_1}{a}\right)\cot k_s h, \tag{42}$$

where h is the half-length of the probe and a is its radius; and Z_A is the self-impedance of the probe when isolated. It is given by (32) or by (29) when losses are negligible.

From (41), it follows that

$$Z_0 = \frac{V_o}{I_t} = Z_L - \frac{I_A}{I_t}(Z_L - Z_A). \tag{43}$$

The current I_A is the current in the probe required to maintain the field $E_z(a_2)$ just outside the shield, if the shield were absent. I_t is the current required to maintain the field $E_z^i(a_1)$ just inside the shield. Since the thickness of the shield is small compared with any radial distance over which E_z could vary significantly in amplitude, it follows that

$$\frac{I_A}{I_t} \sim \frac{E_z(a_2)}{E_z^i(a_1)}. \tag{44}$$

This complex ratio isappr oximately $E_z(a_1)/E_z^i$, obtained when (39) is combined with (17). The substitution of (44) in (43) gives

$$Z_0 = Z_L - \frac{E_z(a_2)}{E_z^i(a_1)}\{Z_L - Z_A\}. \tag{45}$$

This is the final formula for the driving-point impedance of an electric probe symmetrically situated within an imperfectly conducting cylinder.

It is important to observe that the ratio $E_z(a_2)/E_z^i(a_1)$ in (45) is small, since when the probe is driven $E_z(a_2)$, the field on the outside of the shield is small, while the field on the inside of the shield, $E_z^i(a_1)$, is large. On the other hand in the receiving case E_z^i is large and $E_z(a_1)$ is small, thus

$$\left\{\frac{E_z(a_2)}{E_z(a_1)}\right\}=\left\{\frac{E_z(a_1)}{E_z(a_2)}\right\}. \tag{46}$$

Driven Probe Receiving Probe

Since the ratio $E_z(a_2)/E_z^i(a_1)$ is small when the dipole is driven and the cylinder is a moderately good conductor, it follows that

$$Z_0 \approx Z_L. \tag{47}$$

On the other hand when the ratio $E_z(a_2)/E_z^i(a_1) \rightarrow 1$, as when the cylindrical shield is nonexistent,

$$Z_0 \approx Z_A. \tag{48}$$

These relations for limiting values serve as checks on (45).

The equivalent receiving circuit of the probe consists of a generator with an emf equal to $2h_eE_z(a_1)$ driving a circuit consisting of Z_0 in series with the load impedance. The power in the load is readily calculated in terms of the field $E_z(a_1)$. Thus

$$P=|I^2(0)|R_L;\ I(0)=\frac{2h_eE_z(a_1)}{(R_0+R_L)+j(X_0+X_L)}. \tag{49}$$

The impedance of an electric probe is essentially a capacitive reactance since the resistive component is only a fraction of an ohm. Maximum power is obtained in a given load when the probe is tuned to resonance by means of series inductors. If this is done, the full open-circuit, V_{oc} appears across the load, if all losses are neglected. In this case Z_0 does not appear in the equivalent circuit of the probe, and (45) is not needed.

5. Conclusion

A method has been presented for calculating the power in the load of an electric probe encased in a partially conducting cylindrical shield. The analysis has been carried out for sinusoidal signals. If the signal source emits repetitive pulses, it is necessary to make a Fourier analysis of the incoming wave and solve the problem for the fundamental and several harmonic frequencies. The total power in the load impedance is then the sum of the powers absorbed at each frequency considered. Again it is assumed that the circuit dimensions are small in terms of the shortest wavelength of the signal component that contributes significantly to the power in the load.

The field incident upon the cylindrical shield need not be linearly polarized, as assumed in the analysis. An elliptically polarized electric field may be decomposed into two components which are fixed in space and which differ in phase, magnitude, and direction. In particular, the field may be resolved into two mutually perpendicular components which remain stationary in space. One of these components may be chosen parallel to the cylindrical shield, and the other perpendicular to it. Both vary periodically in time, but the latter contributes nothing to the induced voltage; its magnitude and phase are of no significance, and it may be ignored. Accordingly, maximum power is delivered to the load impedance of the probe when the hollow cylinder is oriented in space so that the major axis of the elliptical contour of the electric field is directed parallel to the axis of the cylindrical shield.

6. References

[1] A somewhat different analysis from that presented here for the field inside a conducting tube is given in R. W. P. King, Electromagnetic Engineering, pp. 350 to 359 (McGraw-Hill Book Co., Inc., New York, N.Y., 1945).

[2] R. W. P. King, Theory of Linear Antennas, Ch. IV, Sec. 4 (Harvard University Press, 1956).

[3] More accurate formulas for the effective length and impedance of an electrically short antenna, specifically when $l/a_2=75$ are given by R. W. P. King, Theory of Linear Antennas, supra, p. 496. Specific formulas are $2l_e=0.958l\{1+0.061\,(k_sl)^2\}$; $Z_0^e=18.3\,k_s^2l^2\,(1+0.086\,k_s^2l^2)-j\,\frac{396}{k_sl}\,(1-0.383\,k_s^2l^2)$ ohms.

[4] R. W. P. King, Theory of Linear Antennas, supra, p. 184, eq. (6a).

[5] R. W. P. King, Electromagnetic Engineering, supra, p. 357, eq. (4).

Supplementary References

Louis V. King, Electromagnetic shielding at radio frequencies, Philosophical Magazine **15,** 201 (Feb. 1933).

H. Kaden, Die elektromagnetische Schirmung in der Fernmelde und Hochfrequenztechnik (Springer-Verlag, 1950).

Boulder, Colo.

(Paper 64D3–60)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 3, May–June 1960

Impedance Characteristics of a Uniform Current Loop Having a Spherical Core

Saburo Adachi[1]

(November 12, 1959)

The radiation impedance is derived by the electromotive force method in a convenient form as the sum of the self-radiation impedance of a loop in the free space and an additional term due to the reaction between the loop and the sphere which is proportional to the well-known expansion coefficient of a magnetic-type scattered wave from a sphere in an incident plane wave. The first antiresonance frequency has been given in the form of a universal curve for a very small uniform current loop with core of an arbitrary composition of μ_s and ϵ_s, subject to the condition that the refraction coefficient $N=\sqrt{\mu_s\epsilon_s}$ is extremely large. Some numerical calculations show that high -μ core is desirable for a comparatively lower frequency region, and high -ϵ core is rather desirable in an antiresonance region.

1. Introduction

By use of a sufficiently high refractive index, large scattering or absorption cross sections may be obtained from spheres small in terms of wavelength. For example, it has been shown that the echoing area can be increased to approximately three-quarters of a square wavelength for resonance of the electric or magnetic dipole type, and the required index of refraction can be given as a function of sphere size.

In a small loop antenna the ohmic resistance normally exceeds the radiation resistance. An increase in radiation resistance is highly desirable to improve the radiation efficiency. It can be expected that a high induced voltage, and therefore a high-radiation resistance, can be obtained by encircling a small resonant sphere by such a loop. Of course, it is desired that the radiation reactance be as low as possible. The ratio of the radiation resistance to the radiation reactance is also an important criterion.

Very recently, the most general theoretical analysis for a thin loop with a spherical core has been given by Herman [1][2] and Cruzan [2]. In these analyses, however, it is not easy to derive general relationships between the input impedance and the medium of the spherical core because of the difficulties of the numerical computations.

The purpose of the present paper is to correlate the medium constants of the core with the impedance characteristics under the assumption of a uniform current distribution.[3] In order to derive the radiation impedance, a conventional emf method is applied to the electromagnetic field solution of a uniform loop current in the presence of a sphere which has been treated independently by Tai [3] and Wait [4]. The radiation impedance is given as the sum of the self-radiation impedance of the air loop and an additional term due to the reaction between the loop and the sphere which is proportional to the well-known expansion coefficient of the magnetic-type wave scattered from a sphere in an incident plane wave.

2. Impedance Characteristics of a Uniform Current Loop Having a Spherical Core

The general solution for the electromagnetic field due to a uniform current loop in the presence of a sphere has been given by C. T. Tai. Geometrical configurations are illustrated in figure 1. The radii of the sphere and the loop are the same and are

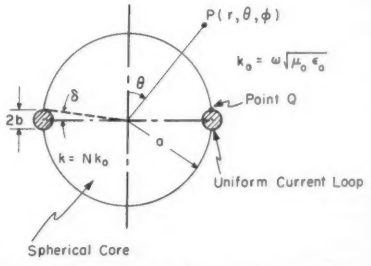


FIGURE 1. *Uniform current loop having a spherical core.*

indicated by a; therefore, the wire of the loop is partially immersed into the core. Wavenumbers of the free space and the medium of the sphere are indicated by k_0 and by $k=Nk_0(N=\sqrt{\epsilon_s\mu_s})$, respectively. As the time dependence, $e^{j\omega t}$ is used. The primary fields due to the uniform loop current, I, are expressed by the following equations:

[1] Antenna Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus 10, Ohio; while author was on leave of absence from Tohoku University, Sendai, Japan (invited paper).

[2] Figures in brackets indicate the literature references at the end of this paper.

[3] The assumption of constancy of current in the loop requires that the loop be fed at a number of points around the loop. In view of the large magnitude of Nka a localized generator would produce a highly nonuniform current.

$$\left.\begin{aligned}\overline{E}^i&=\sum_{n=1}^{\infty} C_n\overline{M}_{\text{eon}}^{(1)}(k_0)\\ \overline{H}^i&=\frac{k_0}{j\omega\mu_0}\sum_{n=1}^{\infty} C_n\overline{N}_{\text{eon}}^{(1)}(k_0)\end{aligned}\right\}\quad r\leq a, \tag{1}$$

and

$$\left.\begin{aligned}\overline{E}^i&=\sum_{n=1}^{\infty} C_n\frac{j_n(\alpha)}{h_n^{(2)}(\alpha)}\overline{M}_{\text{eon}}^{(4)}(k_0)\\ \overline{H}^i&=\frac{k_0}{j\omega\mu_0}\sum_{n=1}^{\infty} C_n\frac{j_n(\alpha)}{h_n^{(2)}(\alpha)}\overline{N}_{\text{eon}}^{(4)}(k_0)\end{aligned}\right\}\quad r\geq a, \tag{2}$$

where $\alpha=k_0a$, and

$$C_n=-\frac{I\omega\mu_0\alpha}{2}\frac{2n+1}{n(n+1)}P_n^1(0)h_n^{(2)}(\alpha). \tag{3}$$

The vector wave functions $\overline{M}_{\text{eon}}$ and $\overline{N}_{\text{eon}}$ are the same with $\overline{m}_{\text{eon}}$ and $\overline{n}_{\text{eon}}$ defined in the text by Stratton [5]. The fields reflected to the external region and the field transmitted into the sphere are, respectively, given by

$$\left.\begin{aligned}\overline{E}^r&=\sum_{n=1}^{\infty} C_nR_n\overline{M}_{\text{eon}}^{(4)}(k_0)\\ \overline{H}^r&=\frac{k_0}{j\omega\mu_0}\sum_{n=1}^{\infty} C_nR_n\overline{N}_{\text{eon}}^{(4)}(k_0)\end{aligned}\right\}\quad r\geq a \tag{4}$$

and

$$\left.\begin{aligned}\overline{E}^t&=\sum_{n=1}^{\infty} C_nT_n\overline{M}_{\text{eon}}^{(1)}(Nk_0)\\ \overline{H}^t&=\frac{k}{j\omega\mu}\sum_{n=1}^{\infty} C_nT_n\overline{N}_{\text{eon}}^{(1)}(Nk_0)\end{aligned}\right\}\quad r\leq a \tag{5}$$

where

$$R_n=\frac{\mu_0 j_n(\alpha)\,[N\alpha j_n(N\alpha)]'-\mu j_n(N\alpha)\,[\alpha j_n(\alpha)]'}{\mu j_n(N\alpha)\,[\alpha h_n^{(2)}(\alpha)]'-\mu_0 h_n^{(2)}(\alpha)\,[N\alpha j_n(N\alpha)]'} \tag{6}$$

$$T_n=\mu\frac{j_n(\alpha)\,[\alpha h_n^{(2)}(\alpha)]'-h_n^{(2)}(\alpha)\,[\alpha j_n(\alpha)]'}{\mu j_n(N\alpha)\,[\alpha h_n^{(2)}(\alpha)]'-\mu_0 h_n^{(2)}(\alpha)\,[N\alpha j_n(N\alpha)]'}. \tag{7}$$

It is very interesting to note that the coefficient R_n is exactly the same as the coefficient associated with the magnetic-type wave scattered by a sphere in an incident plane wave. The coefficients R_n and T_n express the effect of the spherical core on the performance of the loop antenna. Neither the electric field nor the magnetic field (the surface current) are uniform around the wire cross section. If, however, the wire is sufficiently thin, it is reasonable to assume the electric field at the point Q (on the surface of the sphere, and at a distance b, the radius of the wire, from the current filament which is located at the equator of the sphere) as the average value of the electric field around the wire cross section; namely,

$$\begin{aligned}P&\doteq-2\pi a(E_\phi^i+E_\phi^r)^{\substack{r=a\\ \theta=\theta_0}}\int_c(H_t^i+H_t^r)^*ds\\ &=-2\pi aI^*(E_\phi^i+E_\phi^r)^{\substack{r=a\\ \theta=\theta_0}}\end{aligned} \tag{8}$$

where

$$\theta_0=\pi/2-\delta\simeq\pi/2-\frac{b}{a}.$$

Thus, the radiation impedance obtained from the emf method is expressed by

$$Z=\frac{P}{|I|^2}=Z_0+Z_s \tag{9}$$

where

$$Z_0=\pi\eta\alpha^2\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}P_n^1(0)\,P_n^1(\cos\theta_0)j_n(\alpha)h_n^{(2)}(\alpha). \tag{10}$$

$$Z_s=\pi\eta\alpha^2\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}P_n^1(0)P_n^1(\cos\theta_0)R_nh_n^{(2)}(\alpha)^2. \tag{11}$$

In the above equations, Z_0 represents the radiation impedance of a loop without a spherical core, Z_s represents the variation of the radiation impedance due to the reaction between the spherical core and the loop, and η is the intrinsic impedance of free space. The imaginary part of series (10) does not converge rapidly enough when θ_0 approaches $\pi/2$, or the thickness of the wire decreases to zero. By using the asymptotic expression of Legendre and Bessel functions, the radiation reactance X_0 in (10) can be transformed into the following alternative equations:

$$\begin{aligned}X_0=\eta\alpha&\left\{\ln\frac{a}{2b}+\frac{\pi}{4}\frac{b}{a}\right\}\\ &-\pi\eta\alpha^2\sum_{n=1,3}^{\infty}\frac{1}{n+1}\left[\frac{2n+1}{n}P_n^1(0)P_n^1(\cos\theta_0)j_n(\alpha)n_n(\alpha)\right.\\ &\left.+\frac{1}{\pi\alpha}\left\{2\cos\frac{b}{a}(n+1)+\frac{b}{a}\sin\frac{b}{a}(n+1)\right\}\right].\end{aligned} \tag{12}$$

The detailed derivation of the above equation is given in reference [6]. For a small loop, i.e., $\alpha\ll1$, the radiation impedance is approximated by retaining the term $n=1$ only; namely,

$$Z_0 \simeq 20\pi^2\alpha^4 + j60\pi\alpha\left\{2\ln\frac{8a}{b} - 4.40\right\}. \quad (13)$$

As for the constant R_n, many discussions have been made relating to the problem of the scattering of a plane wave by a sphere. These results can be directly applied to the present problem. For the lossy spherical core, i.e., the complex values of N, it is convenient for the computation to transform eq (11) by using the logarithmic derivative functions [7] with respect to spherical Bessel functions into the following form:

$$Z_s = -\pi\eta\alpha^2 \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} P_n^1(0) P_n^1(\cos\theta_0) j_n(\alpha) h_n^{(2)}(\alpha) \times \left\{ \frac{\sigma_n(\alpha) - N\frac{\mu_0}{\mu}\sigma_n(N\alpha)}{\rho_n(\alpha) - N\frac{\mu_0}{\mu}\sigma_n(N\alpha)} \right\} \quad (14)$$

where $\rho_n(\alpha)$ and $\sigma_n(\alpha)$ are the logarithmic derivative functions with respect to spherical Hankel and Bessel functions, respectively, i.e.,

$$\rho_n(x) = \frac{d}{dx}\ln[xh_n^{(2)}(x)]$$

$$\sigma_n(x) = \frac{d}{dx}\ln[xj_n(x)]. \quad (15)$$

The function $\sigma_n(N\alpha)$ of any order with complex arguments $N\alpha$ can be computed by the recurrence formula of the logarithmic derivative functions.

Consider the maximum value of R_n and its condition. A discussion has been given for special cases of pure dielectric ($\mu_s=1$) or magnetic ($\epsilon_s=1$) materials by E. M. Kennaugh. General cases for an arbitrary composition of ϵ_s and μ_s can be discussed in a similar way. When a sphere is sufficiently small in terms of wavelength, $\alpha \ll 1$, only the term R_1 is significant. It is concluded that the maximum absolute value of R_1 is -1 at the condition of

$$\frac{\mu}{\mu_0} j_1(N\alpha)[\alpha n_1(\alpha)]' - n_1(\alpha)[N\alpha j_1(N\alpha)]' = 0. \quad (16)$$

For $\alpha \ll 1$, the above equation is approximately rewritten in the following transcendental equation:

$$\frac{1}{N\alpha} + \frac{N\alpha}{\mu_s - 1} = \cot N\alpha. \quad (17)$$

This equation is graphically solved with respect to $N\alpha$. Figure 2 shows $N\alpha/\pi$ as a function of μ_s. Since α is assumed to be very small compared to unity, N is necessarily very large, i.e., $\sqrt{\mu_s\epsilon_s} \gg 1$. This value gives the first antiresonance frequency of a small uniform-current loop with a lossless spherical core having arbitrary values ϵ_s and μ_s. It should be noted here that the above discussion can be applied to the problem of the plane wave scattering

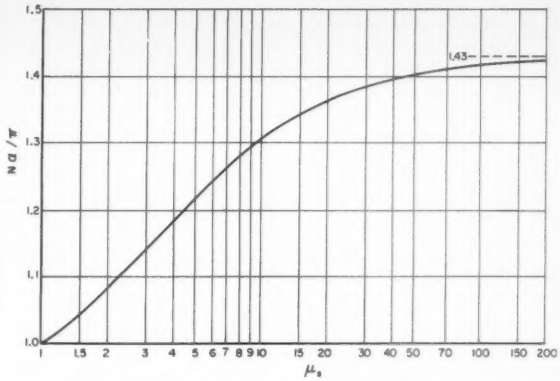


FIGURE 2. *Antiresonance frequency of a uniform current loop having a spherical core with arbitrary constants μ_s and ϵ_s, where $N\alpha = \sqrt{\mu_s\epsilon_s}\, k_0 a = ka$.*

by a small sphere; namely, the curve of figure 2 gives the magnetic-type resonance frequency of the sphere, and also gives the electric-type resonance frequency by replacing μ_s by ϵ_s.

The maximum radiation resistance at the antiresonance point is given by

$$R = R_0 + R_s = \frac{3}{2}\pi\eta\alpha^2 n_1(\alpha)^2 \simeq \frac{3}{2}\pi\eta\frac{1}{\alpha^2}, \ (\alpha \ll 1). \quad (18)$$

The maximum resistance increases in the order of α^{-2} with the decrease of α. It should be noted that in an actual one-point-fed loop, unless infinitesimally small, the frequency of its antiresonance deviates from that predicted by figure 2 due to the nonuniform current distribution. This situation is illustrated in table 1.

The diameter of the antenna in terms of a wavelength in the medium at an antiresonance is found to vary between 0.159 and 1.43 by the constants of a medium. The results in figure 2 give the limiting case of an infinitesimally small loop, i.e., $\mu_s\epsilon_s \to \infty$. Similarly, the expression for Z in eq (9) is correct only for an infinitesimally small loop, if it is fed at one point. The radiation impedance Z has been calculated for the two cases $\epsilon_s=100$, $\mu_s=1$, and $\mu_s=100$, $\epsilon_s=1$. For simplicity, the core is assumed to be lossless. Figure 3 and 4 show the radiation resistances and the radiation reactances. It must be noted here also, as mentioned before, that these figures do not give quantitative results for a one-point-fed loop be-

TABLE 1. *Relation between medium of core and dimensions at antiresonance*

Medium of core	α	$2a$
Air	0.5	0.159λ
*$\mu_s = 10.8 - j0.24$, $\epsilon_s = 9.5 - j0.077$	.1535	.496λ
$(\mu_s\epsilon_s) \to \infty$	$\frac{\pi}{\sqrt{\mu_s\epsilon_s}}$ (1.00~1.43)	(1.00~1.43)λ

*From ref [1] by Julius Herman, λ is the wavelength in the medium of the core.

$$\left.\begin{aligned}\overline{E}^i&=\sum_{n=1}^{\infty} C_n \overline{M}_{\text{eon}}^{(1)}(k_0)\\ \overline{H}^i&=\frac{k_0}{j\omega\mu_0}\sum_{n=1}^{\infty} C_n \overline{N}_{\text{eon}}^{(1)}(k_0)\end{aligned}\right\} \quad r\leq a, \tag{1}$$

and

$$\left.\begin{aligned}\overline{E}^i&=\sum_{n=1}^{\infty} C_n \frac{j_n(\alpha)}{h_n^{(2)}(\alpha)}\overline{M}_{\text{eon}}^{(4)}(k_0)\\ \overline{H}^i&=\frac{k_0}{j\omega\mu_0}\sum_{n=1}^{\infty} C_n \frac{j_n(\alpha)}{h_n^{(2)}(\alpha)}\overline{N}_{\text{eon}}^{(4)}(k_0)\end{aligned}\right\} \quad r\geq a, \tag{2}$$

where $\alpha=k_0a$, and

$$C_n=-\frac{I\omega\mu_0\alpha}{2}\frac{2n+1}{n(n+1)}P_n^1(0)h_n^{(2)}(\alpha). \tag{3}$$

The vector wave functions $\overline{M}_{\text{eon}}$ and $\overline{N}_{\text{eon}}$ are the same with $\overline{m}_{\text{eon}}$ and $\overline{n}_{\text{eon}}$ defined in the text by Stratton [5]. The fields reflected to the external region and the field transmitted into the sphere are, respectively, given by

$$\left.\begin{aligned}\overline{E}^r&=\sum_{n=1}^{\infty} C_nR_n\overline{M}_{\text{eon}}^{(4)}(k_0)\\ \overline{H}^r&=\frac{k_0}{j\omega\mu_0}\sum_{n=1}^{\infty} C_nR_n\overline{N}_{\text{eon}}^{(4)}(k_0)\end{aligned}\right\} \quad r\geq a \tag{4}$$

and

$$\left.\begin{aligned}\overline{E}^t&=\sum_{n=1}^{\infty} C_nT_n\overline{M}_{\text{eon}}^{(1)}(Nk_0)\\ \overline{H}^t&=\frac{k}{j\omega\mu}\sum_{n=1}^{\infty} C_nT_n\overline{N}_{\text{eon}}^{(1)}(Nk_0)\end{aligned}\right\} \quad r\leq a \tag{5}$$

where

$$R_n=\frac{\mu_0 j_n(\alpha)\,[N\alpha j_n(N\alpha)]'-\mu j_n(N\alpha)\,[\alpha j_n(\alpha)]'}{\mu j_n(N\alpha)\,[\alpha h_n^{(2)}(\alpha)]'-\mu_0 h_n^{(2)}(\alpha)\,[N\alpha j_n(N\alpha)]'} \tag{6}$$

$$T_n=\mu\frac{j_n(\alpha)\,[\alpha h_n^{(2)}(\alpha)]'-h_n^{(2)}(\alpha)\,[\alpha j_n(\alpha)]'}{\mu j_n(N\alpha)\,[\alpha h_n^{(2)}(\alpha)]'-\mu_0 h_n^{(2)}(\alpha)\,[N\alpha j_n(N\alpha)]'}. \tag{7}$$

It is very interesting to note that the coefficient R_n is exactly the same as the coefficient associated with the magnetic-type wave scattered by a sphere in an incident plane wave. The coefficients R_n and T_n express the effect of the spherical core on the performance of the loop antenna. Neither the electric field nor the magnetic field (the surface current) are uniform around the wire cross section. If, however, the wire is sufficiently thin, it is reasonable to assume the electric field at the point Q (on the surface of the sphere, and at a distance b, the radius of the wire, from the current filament which is located at the equator of the sphere) as the average value of the electric field around the wire cross section; namely,

$$\begin{aligned}P&\doteq-2\pi a(E_\phi^i+E_\phi^r)_{\theta=\theta_0}^{r=a}\cdot\int_c(H_l^i+H_l^r)^*ds\\ &=-2\pi aI^*(E_\phi^i+E_\phi^r)_{\theta=\theta_0}^{r=a}\end{aligned} \tag{8}$$

where

$$\theta_0=\pi/2-\delta\simeq\pi/2-\frac{b}{a}.$$

Thus, the radiation impedance obtained from the emf method is expressed by

$$Z=\frac{P}{|I|^2}=Z_0+Z_s \tag{9}$$

where

$$Z_0=\pi\eta\alpha^2\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}P_n^1(0)P_n^1(\cos\theta_0)j_n(\alpha)h_n^{(2)}(\alpha). \tag{10}$$

$$Z_s=\pi\eta\alpha^2\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}P_n^1(0)P_n^1(\cos\theta_0)R_nh_n^{(2)}(\alpha)^2. \tag{11}$$

In the above equations, Z_0 represents the radiation impedance of a loop without a spherical core, Z_s represents the variation of the radiation impedance due to the reaction between the spherical core and the loop, and η is the intrinsic impedance of free space. The imaginary part of series (10) does not converge rapidly enough when θ_0 approaches $\pi/2$, or the thickness of the wire decreases to zero. By using the asymptotic expression of Legendre and Bessel functions, the radiation reactance X_0 in (10) can be transformed into the following alternative equations:

$$\begin{aligned}X_0=\eta\alpha&\left\{\ln\frac{a}{2b}+\frac{\pi}{4}\frac{b}{a}\right\}\\ &-\pi\eta\alpha^2\sum_{n=1,3}^{\infty}\frac{1}{n+1}\left[\frac{2n+1}{n}P_n^1(0)P_n^1(\cos\theta_0)j_n(\alpha)n_n(\alpha)\right.\\ &\left.+\frac{1}{\pi\alpha}\left\{2\cos\frac{b}{a}(n+1)+\frac{b}{a}\sin\frac{b}{a}(n+1)\right\}\right].\end{aligned} \tag{12}$$

The detailed derivation of the above equation is given in reference [6]. For a small loop, i.e., $\alpha\ll1$, the radiation impedance is approximated by retaining the term $n=1$ only; namely,

$$Z_0 \simeq 20\pi^2\alpha^4 + j60\pi\alpha\left\{2\ln\frac{8a}{b} - 4.40\right\}. \quad (13)$$

As for the constant R_n, many discussions have been made relating to the problem of the scattering of a plane wave by a sphere. These results can be directly applied to the present problem. For the lossy spherical core, i.e., the complex values of N, it is convenient for the computation to transform eq (11) by using the logarithmic derivative functions [7] with respect to spherical Bessel functions into the following form:

$$Z_s = -\pi\eta\alpha^2 \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} P_n^1(0) P_n^1(\cos\theta_0) j_n(\alpha) h_n^{(2)}(\alpha) \times \left\{ \frac{\sigma_n(\alpha) - N\frac{\mu_0}{\mu}\sigma_n(N\alpha)}{\rho_n(\alpha) - N\frac{\mu_0}{\mu}\sigma_n(N\alpha)} \right\} \quad (14)$$

where $\rho_n(\alpha)$ and $\sigma_n(\alpha)$ are the logarithmic derivative functions with respect to spherical Hankel and Bessel functions, respectively, i.e.,

$$\rho_n(x) = \frac{d}{dx}\ln[xh_n^{(2)}(x)]$$

$$\sigma_n(x) = \frac{d}{dx}\ln[xj_n(x)]. \quad (15)$$

The function $\sigma_n(N\alpha)$ of any order with complex arguments $N\alpha$ can be computed by the recurrence formula of the logarithmic derivative functions.

Consider the maximum value of R_n and its condition. A discussion has been given for special cases of pure dielectric ($\mu_s=1$) or magnetic ($\epsilon_s=1$) materials by E. M. Kennaugh. General cases for an arbitrary composition of ϵ_s and μ_s can be discussed in a similar way. When a sphere is sufficiently small in terms of wavelength, $\alpha \ll 1$, only the term R_1 is significant. It is concluded that the maximum absolute value of R_1 is -1 at the condition of

$$\frac{\mu}{\mu_0} j_1(N\alpha)[\alpha n_1(\alpha)]' - n_1(\alpha)[N\alpha j_1(N\alpha)]' = 0. \quad (16)$$

For $\alpha \ll 1$, the above equation is approximately rewritten in the following transcendental equation:

$$\frac{1}{N\alpha} + \frac{N\alpha}{\mu_s - 1} = \cot N\alpha. \quad (17)$$

This equation is graphically solved with respect to $N\alpha$. Figure 2 shows $N\alpha/\pi$ as a function of μ_s. Since α is assumed to be very small compared to unity, N is necessarily very large, i.e., $\sqrt{\mu_s\epsilon_s} \gg 1$. This value gives the first antiresonance frequency of a small uniform-current loop with a lossless spherical core having arbitrary values ϵ_s and μ_s. It should be noted here that the above discussion can be applied to the problem of the plane wave scattering

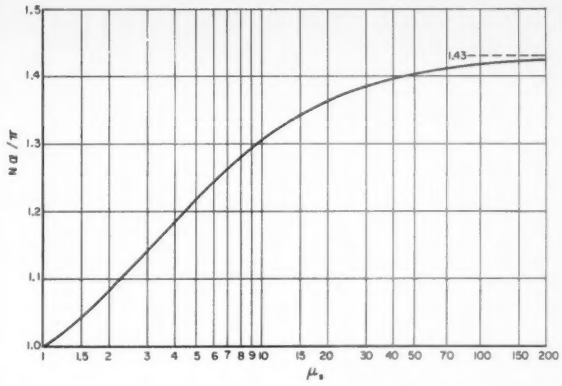


FIGURE 2. *Antiresonance frequency of a uniform current loop having a spherical core with arbitrary constants μ_s and ϵ_s, where $N\alpha = \sqrt{\mu_s\epsilon_s}\, k_0 a = ka$.*

by a small sphere; namely, the curve of figure 2 gives the magnetic-type resonance frequency of the sphere, and also gives the electric-type resonance frequency by replacing μ_s by ϵ_s.

The maximum radiation resistance at the antiresonance point is given by

$$R = R_0 + R_s = \frac{3}{2}\pi\eta\alpha^2 n_1(\alpha)^2 \simeq \frac{3}{2}\pi\eta\frac{1}{\alpha^2}, \ (\alpha \ll 1). \quad (18)$$

The maximum resistance increases in the order of α^{-2} with the decrease of α. It should be noted that in an actual one-point-fed loop, unless infinitesimally small, the frequency of its antiresonance deviates from that predicted by figure 2 due to the nonuniform current distribution. This situation is illustrated in table 1.

The diameter of the antenna in terms of a wavelength in the medium at an antiresonance is found to vary between 0.159 and 1.43 by the constants of a medium. The results in figure 2 give the limiting case of an infinitesimally small loop, i.e., $\mu_s\epsilon_s \to \infty$. Similarly, the expression for Z in eq (9) is correct only for an infinitesimally small loop, if it is fed at one point. The radiation impedance Z has been calculated for the two cases $\epsilon_s=100$, $\mu_s=1$, and $\mu_s=100$, $\epsilon_s=1$. For simplicity, the core is assumed to be lossless. Figure 3 and 4 show the radiation resistances and the radiation reactances. It must be noted here also, as mentioned before, that these figures do not give quantitative results for a one-point-fed loop be-

TABLE 1. *Relation between medium of core and dimensions at antiresonance*

Medium of core	α	$2a$
Air	0.5	0.159λ
*$\mu_s = 10.8 - j0.24$, $\epsilon_s = 9.5 - j0.077$	.1535	.496λ
$(\mu_s\epsilon_s) \to \infty$	$\frac{\pi}{\sqrt{\mu_s\epsilon_s}}$ (1.00~1.43)	(1.00~1.43)λ

*From ref [1] by Julius Herman, λ is the wavelength in the medium of the core.

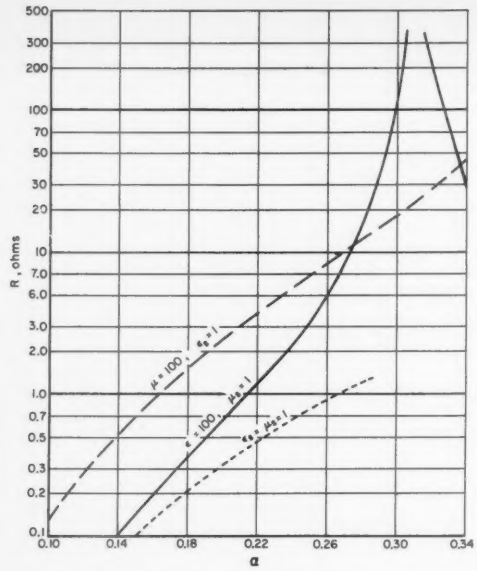


FIGURE 3. *Radiation resistances of uniform current loops having air, dielectric, and magnetic cores* (N=1, 10; a/b=60).

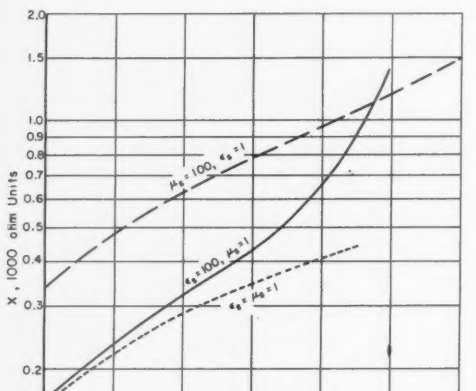


FIGURE 4. *Radiation reactances of uniform current loops having air, dielectric, and magnetic cores* (N=1, 10; a/b=60.)

cause of its nonuniform current. As for the case of $\epsilon_s=100$, the antiresonance occurs at $\alpha=\pi/10$, and the curves approach asymptotically to those of an air loop with the decrease of ρ. The μ-core is effective in order to increase the radiation resistance in the lower frequency region. The radiation impedance for an arbitrary composition of μ and ϵ ought to show a behavior intermediate between these two extreme cases. In figure 5 the ratios of the radiation resistances to the radiation reactances are plotted. From these figures it can be concluded that high-μ core is

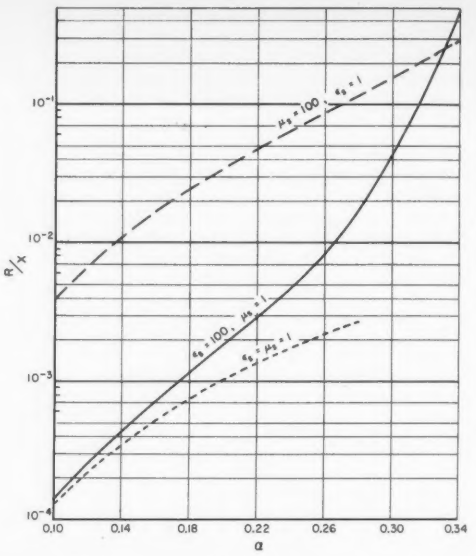


FIGURE 5. *Ratios of radiation resistances to radiation reactances of uniform current loops having air, dielectric, and magnetic cores* (N=1, 10; a/b=60).

desirable for a comparatively low-frequency region, and high-ϵ core is rather desirable in order to use it in a antiresonant frequency region. This can be said also from the viewpoint of medium losses, that is, losses of usual magnetic materials increase more rapidly with the frequency than do the losses of dielectric materials.

3. Conclusion

The impedance characteristics of a uniform current loop (see footnote 3) with a spherical core of an arbitrary composition of μ and ϵ have been theoretically investigated. The radiation impedance is derived by the emf method in a convenient form as the sum of the self-radiation impedance of an air loop and an additional term due to the reaction between the loop and the sphere which is proportional to the well-known expansion coefficient of a magnetic-type scattered wave from a sphere in an incident plane wave. The first antiresonance frequency has been given in the form of a universal curve for a very small uniform current loop with a core of an arbitrary composition of μ_s and ϵ_s, subject to the condition that the refraction coefficient $N=\sqrt{\mu_s\epsilon_s}$ is extremely large. It should be noted that the above results can be directly applied to the first magnetic-type resonance frequency of a sphere in an incident plane wave. To show quantitative impedance characteristics of a loop with an arbitrary composite core, the radiation impedances for two cases, namely $\epsilon_s=100$, $\mu_s=1$, and $\mu_s=100$, $\epsilon_s=1$, have been calculated and compared with an air loop. It is concluded that high-μ core is desirable for a comparatively lower frequency region, and high-ϵ core is rather desirable in an antiresonance frequency region.

The author thanks E. M. Kennaugh for suggesting this investigation as well as for guidance in the course of the work, and also Dr. R. G. Kouyoumjian for his valuable discussions.

4. References

[1] J. Herman, Thin wire loop and thin biconical antennas in finite spherical media, Ph. D. Dissertation, University of Maryland, Publication No. 25, 331 (1957); Further extensions of loop antenna theory, TR–756 (July 1957), Diamond Ordnance Fuze Laboratory.

[2] O. R. Cruzan, Radiation properties of a spherical ferrite antenna, TR–387 (Oct. 1956), VHF ferrite antenna radiation properties, TR–516 (Aug. 1957), Diamond Ordnance Fuze Laboratory.

[3] C. T. Tai, Radiation from a uniform circular loop antenna in the presence of a sphere, Stanford Research Instit., Tech. Rept. No. 32 (1952).

[4] J. R. Wait, Electromagnetic coupling between a circular loop and a conducting sphere, Geophysics **18,** 971 (1953); see also Ph. D. Thesis, University of Toronto (March 1951). The loop inside a spherical cavity in a dissipative medium has also been considered by J. R. Wait, Insulated loop antenna immersed in a conducting medium, J. Research NBS **59,** 133 (1957) RP2781.

[5] J. A. Stratton, Electromagnetic Theory (McGraw-Hill Book Co., Inc., New York, N.Y., 1941).

[6] S. Adachi, Impedance characteristics of a uniform current loop having a spherical core, The Ohio State University Antenna Lab. Tech. Rept., 662 (1959).

[7] A. L. Aden, Electromagnetic scattering from sphere with sizes comparable to the wavelength, J. Appl. Phys. **22,** 601 (1951).

BOULDER, COLO. (Paper 64D3–61)

Publication of the National Bureau of Standards*

Selected Abstracts

The effect of multipath distortion on the choice of operating frequencies for high-frequency communication circuits, D. K. Bailey, *IRE Trans. Ant. Prop.* **AP–7,** *397 (1959).*

Harmful multipath distortion on high-frequency facsimile services and telegraphic services operating at high speeds occurs when the received signal is composed of two or more components arriving by different modes over the same great-circle path with comparable intensities, but having travel times which differ by an amount equal to an appreciable fraction of the duration of a signal element. The dependence of multipath distortion on the relationship of the operating frequency to the MUF is discussed and a new term, the multipath reduction factor (MRF), is introduced which permits calculation in terms of the MUF of the lowest frequency which can be used to provide a specified measure of protection against multipath distortion. The MRF has a marked path-length dependence and is calculated as a function of path length for representative values of the other parameters involved by making use of an ionospheric model. It is then shown how the MRF can be used in connection with world-wide MUF prediction material to determine the minimum number of frequencies which must be assigned to a high-frequency communication service of continuous availability operating at high speed. Some comparisons with observations are discussed, and finally conclusions are drawn concerning manner of operation and choice of operating frequencies to reduce or to eliminate harmful multipath distortion.

Reception of space diversity transmitters—Observations over long-distance path to evaluate the usefulness of the system, J. W. Koch, *Wireless World* **65,** *512 (1959).*

The British Broadcasting Corporation recently provided a series of transmissions to test the effectiveness of transmitter space-diversity on reception at distant points. Both very widely-spaced and relatively closely-spaced transmitters were used in the tests. The transmissions were at a frequency of 9,510 kc/s, directed towards the east coast of the United States. Transmitting conditions were switched at intervals of approximately fifteen minutes during the testing period each day.

Observations of the received signals were made at the National Bureau of Standards Laboratories in Boulder, Colorado, from November 3, 1958 to November 14, 1958. The recordings obtained during these observations have been analyzed for fading characteristics and intelligibility, and the results of the analysis are given in this article.

Mismatch errors in cascade-connected variable attenuators, G. E. Shafer and A. Y. Rumfelt, *IRE Trans. Microw. Theory Tech.* **MMT–7,** *447 (1959).*

The treatment of mismatch errors is extended to cover variable attenuators cascade-connected in a system which is not free from reflections. The method of analysis is applicable to any number of cascaded attenuators, but only the analysis of two and three variable attenuators in cascade is presented. Graphs are given to aid in estimating the limits of mismatch error.

In an example, which is considered representative of rigid rectangular waveguide systems, the limits of error are: for two attenuators in cascade, 0.19 db in a 3-db measurement, and 0.17 db in a 40-db measurement; and for three attenuators in cascade, 0.25 db in a 40-db measurement, and 0.23 db in a 75-db measurement.

Use of Chebychev polynomials in thin film computations, K. D. Mielenz, *J. Research NBS* **63A,** *No.* 3, *297 (1959).*

From Herpin's expression for the mth power of a multilayer matrix, very simple closed formulas are derived for the matrices and optical constants of any multilayer with a periodic structure.

According to Epstein's theorem, any symmetrical multilayer is equivalent to a fictitious monolayer. A simple expression for the equivalent index and thickness of this monolayer is deduced for the case of a periodic and symmetrical sequence of equally thick films.

As compared to any other method of numerical computation, the suggested formulation provides a considerable saving of time and work. In a numerical example, this saving amounts to about 80 percent.

Recent experimental evidence favouring the $\rho K_1(\rho)$ correlation function for describing the turbulence of refractivity in the troposphere and stratosphere, K. A. Norton, *J. Atmospheric and Terrest. Phys.*, **15,** *206 (1959).*

Experimental evidence of three different kinds is given which shows that the variations of the refractive index in space at a fixed time may be described with a useful degree of accuracy by the correlation function $\rho K_1(\rho)$ where K_1 is the modified Bessel function of the second kind and $\rho = rl_0$ is a normalized distance between two points in the atmosphere. The three kinds of evidence are (a) direct measurements of the variations of n with time at a fixed location as made with a refractometer, (b) measurements of the variations with time of

the phase of a radio wave received over a line-of-sight path, and (c) measurements of the received power of radio waves scattered in the forward direction to a receiving point beyond the radio horizon of the transmitters.

Other NBS Publications

Journal of Research, Vol. 64A, No. 2, March–April 1960. 70 cents.

A carbon-14 beta-ray standard, benzoic acid-7-C^{14} in toluene, for liquid scintillation counters. W. F. Marlow and R. W. Medlock.

A comparison of experimental and theoretical relations between Young's modulus and the flexural and longitudinal resonance frequencies of uniform bars. S. Spinner, T. W. Reichard, and W. E. Tefft.

Determination of copolymer composition by combustion analysis for carbon and hydrogen. Lawrence A. Wood, Irving Madorsky, and Rolf A. Paulson.

Some effects of aging on the surface area of portland cement paste. C. M. Hunt, L. A. Tomes, and R. L. Blaine.

Conformations of the pyranoid sugars. I. Classification of conformers. H. S. Isbell and R. S. Tipson.

Tritium-labeled compounds III. Aldoses-*1-t*. H. S. Isbell, H. L. Frush, N. B. Holt, and J. D. Moyer.

Determination of aluminum in precipitation hardening stainless steel and high temperature alloys. Lawrence A. Machlan, John L. Hague, and Edward J. Meros.

CRPL exponential reference atmosphere, B. R. Bean and G. D. Thayer. NBS Monograph 4 (1959) 45 cents.

Preservation of documents by lamination, W. K. Wilson and B. W. Forshee. NBS Monograph 5 (1959) 20 cents.

X-ray attenuation coefficients from 10 kev to 100 Mev, R. T. McGinnies. NBS Circ. 583 Supplement (1959) 15 cents.

Screw-thread standards for federal services 1957. Amends in part H28 (1944) (and in part its 1950 Supplement). NBS Handb. H28 (1957)—Part II (1959) 75 cents.

Distribution of mail by destination at the San Francisco, Los Angeles, and Baltimore Post Offices, N. C. Severo and A. E. Newman. NBS Tech. Note 27 (PB151386) (1959) $1.50.

A history of vertical-incidence ionosphere sounding at the National Bureau of Standards, S. C. Gladden. NBS Tech. Note 28 (PB151387) (1959) $2.00.

Photographic dosimetry at total exposure levels below 20 mr, M. Ehrlich and W. L. McLaughlin. NBS Tech. Note 29 (PB151388) (1959) 50 cents.

Resistance diode bridge circuit for temperature control, L. H. Bennett and V. M. Johnson. NBS Tech. Note 34 (PB151393) (1959) 50 cents.

Spectral study of a visible, short-duration afterglow in nitrogen, B. E. Beale, Jr., and H. P. Broida, J. Chem. Phys. **31,** No. 4, 1030 (1959).

Atomic electron affinities, F. Rohrlich, Nature **183,** 244 (1959).

Cavity resonators for dielectric spectroscopy of compressed gases, H. E. Bussey and G. Birnbaum, Rev. Sci. Instr. **30,** 800 (1959).

A discussion of federal specifications GS–X–620 and L–F–310 for dental X-ray apparatus and dental X-ray film, G. C. Paffenbarger, A. F. Forziati, and M. P. Kumpula, J. Am. Dental Assoc. **59,** 472 (1959).

Dynamic stability of frozen radicals. II. The formal theory of the model, J. L. Jackson, J. Chem. Phys. **31,** No. 3, 722 (1959).

Method of cooling head-on photomultipliers, G. C. Harman, Rev. Sci. Instr. **30,** No. 8, 742 (1959).

On artificial geomagnetic and ionospheric storms associated with high-altitude explosions, S. Matsushita, J. Geophys. Research **6,** 1149 (1959).

An intermittent-action camera with absolute time calibration, R. G. Hefley, R. H. Doherty, and E. L. Berger, IRE Wescon Conv. Record **3,** Pt. 6, 129 (1959).

Nonresonant microwaves absorption and electric dipole moment of NO in the gaseous state, A. A. Maryott and S. J. Kryder, J. Chem. Phys. **31,** No. 3, 617 (1959).

The earth and its environment, S. Chapman, Proc. IRE **47,** No. 2, 137 (1959).

Bremsstrahlung cross-section formulas and related data, H. W. Koch and J. W. Motz, Rev. Modern Phys. **31,** No. 4, 920 (1959).

Short-time stability of a quartz-crystal oscillator as measured with an ammonia maser, A. H. Morgan and J. A. Barnes, Letter Proc. IRE **47,** 1782 (1959).

Electromagnetic radiation from cylindrical structures, J. R. Wait (Pergamon Press, New York, N.Y., 1959).

Exploratory study, by low temperature X-ray diffraction techniques, of diborane and the products of a microwave discharge in diborane, L. H. Bolz, F. A. Mauer, and H. S. Peiser, J. Chem. Phys. **31,** No. 4, 1005 (1959).

Some clinical applications of research findings in dental materials, G. C. Paffenbarger, Ohio Dental J. **33,** No. 3, 218 (1959).

Liquid helium cryostat with an integral superconducting resonator, E. Maxwell and F. A. Schmidt, Suppl. au Bull. Inst. Intern. du Froid, Comm. 1, Delft, Holland, Annexe 1958–1, p. 95 (1958).

Sporadic *E* observed on VHF oblique incidence circuits, E. K. Smith, in North Atlantic Treaty Organization. Sporadic *E* ionization; Ionospheric Research Meeting, AGARD Avionics Panel, Cambridge, England, Sept. 1958 (AGARDograph 34), p. 129 (1958).

Low-temperature transport properties of copper and its dilute alloys: Pure copper, annealed and cold-drawn, R. L. Powell, H. M. Roder, and W. J. Hall, Phys. Rev. **115,** 314 (1959).

The double bond isomerization of olefins by hydrogen atoms at −195°, M. D. Scheer and R. Klein, J. Phys. Chem. **63,** 1517 (1959).

**Publications for which a price is indicated (except for NBS Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (Order by PB number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Relation of Turbulence Theory to Ionospheric Forward Scatter Propagation Experiments[1]

Albert D. Wheelon

(October 28, 1959; revised February 4, 1960)

This paper attempts to relate turbulence theories to radio measurements on vhf ionospheric forward scatter circuits. To this end, the single scattering description of the electromagnetic response of electron density irregularities and the corresponding transmission expression are evaluated. Statistical distributions of signal levels are found to agree with a scattering model. The several theories for turbulent mixing of the electron density are then summarized. A turbulence mixing model is compared favorably with experimental data on absolute signal levels and their diurnal and seasonal variations. Scattering heights in the ionosphere responsible for the signals are also consistent with these theories. Frequency and distance dependence scaling laws are compared briefly with the data. The scatter signal behavior during sudden ionospheric disturbances is also explained. No attempt is made to compare meteoric and turbulence scatter contributions to the measured quantities in this paper. Approximately thirty references are given.

1. Introduction

The chance observation of unexplicably high field strengths beyond the optical horizon at microwave frequencies [1][2] gave the first hint of "scatter propagation." Subsequent experiments showed that the received signal was weak but significantly greater than that predicted by round earth diffraction theory [2]. The field strength decreased slowly with distance and was observed to fade about its mean value several times a second. These observations suggested a statistical theory of scattering from turbulent fluctuations of the tropospheric refractive index [3, 4]; a theory which has been developed and fitted to increasingly refined experiments [5].

It was inferred from the fading of ionospherically reflected skywave signals that similar refractive irregularities must also exist in the lower region (E and D layers) of the ionosphere. This evidence plus the reality of tropospheric scatter suggested that a similar propagation mode might be sustained at frequencies above the maximum usable frequency (MUF) by scattering from turbulent concentrations of free electrons in the ionosphere. Such a mechanism was immediately observed at vhf for ranges extending to 1,500 km [6], thus providing an important new mode of radio propagation.

The phenomenon of ionospheric scatter propagation has also evinced a growing scientific interest. The basic theoretical problem combines electromagnetic theory, probabilistic descriptions of random signals, and turbulence theory in an essential mixture. On the experimental side, radio waves represent an accurate, reliable tool for studying turbulent conditions at high Reynolds numbers in the ionosphere.

This paper is presented as a linkage between the turbulence theories and the radio measurements. It attempts to point up the important assumptions and questionable approximations in the present transmission expressions by way of emphasizing the accuracy of this linkage. At the same time, brief accounts of the experimental data and theoretical models are given in the interest of acquainting the radio physicists and turbulence theoreticians with the implications of the basic results in the opposite field. The present paper is preliminary in that it anticipates some results of a thoroughgoing comparison of theory and experiment for ionospheric scatter propagation, which is to be prepared shortly. This paper does not consider the contribution of small meteors to the background scatter signal.

2. Statistical Behavior of the Signal

Random fading of the received signal is one of the principal reasons for identifying ionospheric scatter propagation with electromagnetic scattering from turbulent irregularities. A typical time history of the amplitude R of the complex voltage $x+iy$ induced in the receiver by the incident electromagnetic field is shown in figure 1. Statistical analysis of actual records show that the probability that the amplitude exceeds a given level is approximately described by a Rayleigh distribution [7, 8]. This suggests that the process $R(t)$ is the vector amplitude of two orthogonal (phase quadrature) Gaussian random processes. On the other hand, if the received signal were produced by a large number of independent scattering contributions, one would expect the signal components to be distributed in a Gaussian manner by the Central Limit Theorem [9]. This line of reasoning indicates that scattering theory provides a reasonable (though not unique) explanation for the statistical distribution of the experimental data.

[1] Contribution from Space Technology Laboratories, Inc., Los Angeles 45, Calif. An abbreviated version of this paper was published in J. Geophys. Research 64, 223 (Dec. 1959).

[2] Figures in brackets indicate the literature references at the end of this paper.

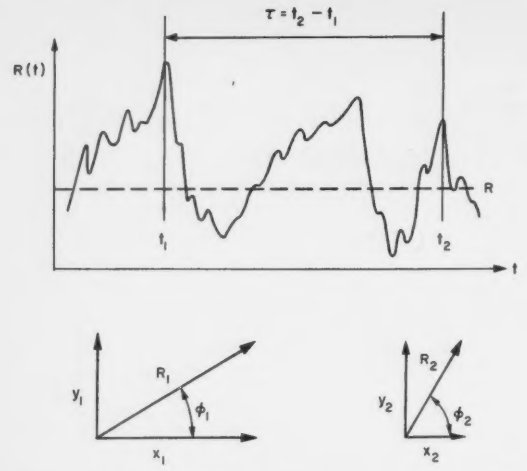


FIGURE 1. *Typical amplitude versus time record of ionospheric scatter propagation signal amplitude together with corresponding time—displaced vector voltage diagrams which are assumed to describe the process.*

To be more precise about these ideas, it is important to set down their corresponding analytical expressions. The joint probability distribution for the Gaussian random processes $x(t)$ and $y(t)$ is,

$$P(x,y)\,dx\,dy=\frac{dx\,dy}{2\pi\sigma^2\sqrt{1-\rho^2}}\exp-\left[\frac{x^2+y^2-2\rho xy}{2\sigma^2(1-\rho^2)}\right] \quad (2.1)$$

where

$$\langle x^2\rangle=\langle y^2\rangle=\sigma^2$$

$$\langle xy\rangle=\sigma^2\rho.$$

When eq (2.1) is transformed to the polar coordinates (R, ϕ) of figure 1, and the phase averaged out, the amplitude distribution becomes:

$$P(R)R\,dR=\frac{R\,dR}{\sigma^2\sqrt{1-\rho^2}}\,I_o\left(\frac{R^2}{2\sigma^2}\frac{\rho}{1-\rho^2}\right)\exp-\left[\frac{R^2}{2\sigma^2(1-\rho^2)}\right]. \quad (2.2)$$

This expression reduces to the Rayleigh distribution if the component correlation ρ is assumed to be zero. However, this assumption is probably not justified for electromagnetic scattering—as it is for electrical shot noise—and the more general expression (2.2) may explain interesting anomalies in the experimental envelope distribution data [10]. The data generally seem to emphasize larger values of R (relative to the mean) more than the Rayleigh distribution predicts, with corresponding reductions in the occurrence of small values. This behavior could also be explained by: (1) Admixture of a steady signal component—perhaps due to partial ionospheric reflections above the MUF, or (2) trends in the average signal, which would produce larger apparent deviations from the mean level when viewed over a finite data sample. On the other hand, such anomalies may represent the effect of meteor-reflected signals.

One can also discuss the joint amplitude distribution for two amplitude processes R_1 and R_2 [5]. These could refer to signals received at two adjacent aerials, such as are often used in diversity receiving systems. R_1 and R_2 could also refer to two time-displaced records of the same circuit signals, as shown in figure 1. Dual process distributions introduce additional statistical parameters μ, which represent the (time, space, etc.) correlations between the component processes x_1, y_1, x_2, y_2. The purely probabilistic predictions of such distributions have been confirmed on tropospheric scatter circuits [11] but have yet to be studied carefully for ionospheric scatter.

In general, one can say that the statistical predictions of scattering theory are consistent with the experimental data. This agreement is neither exact nor unique, but it is satisfactory. The real problem in scatter theory is to *predict* successfully the statistical parameters, σ, ρ, μ, etc., as functions of the variables which are experimentally controlled: radio frequency, path distance, time of day and season, etc. No theoretical estimates for the correlation coefficients ρ and μ have been published, with the exception of certain calculations for tropospheric scatter [5], which may or may not be applicable to ionospheric circuits. Research to date has been focused exclusively on the average power, which according to eq (2.2) is

$$\langle R^2\rangle=2\sigma^2. \quad (2.3)$$

3. Electromagnetic Scattering by Turbulent Irregularities

The basic wave equation which discribes radio-wave propagation through an ionized plasma of electron density N_0 with stochastic fluctuations $\boldsymbol{\delta N}$ added thereto is:

$$[\nabla^2+k^2-r_0(N_0+\boldsymbol{\delta N}]\,\vec{E}=0, \quad (3.1)$$

where $k=2\pi/\lambda$ is the free space wavenumber and $r_0=2.8\times10^{-13}$ cm the classical electron radius. The mean electron density N_0 in (3.1) is usually omitted, since refractive effects have been found to be negligible [12]. In establishing expression (3.1), one assumes that: (1) The stochastic variations $\boldsymbol{\delta N}(\vec{r},t)$ rearrange themselves in a time which is long compared to the period of the radio oscillations, and (2) that gradients of $\boldsymbol{\delta N}$ are negligible over a wavelength. The latter assumption is clearly the more questionable.

It is not possible to solve eq (3.1) exactly, because $\boldsymbol{\delta N}(r,t)$ is an unknown stochastic function of position and time. Heretofore, the Born approximation has

been used exclusively to describe single scattering by such fluctuations. This is essentially a three-dimensional iterative solution of the wave eq (3.1), which gives the scattered field strength at the receiver R in terms of the unperturbed wave $E_0(r)$.

$$E_s(R)=r_o\int_V \mathrm{d}^3r\; G\,(R,r)\;\boldsymbol{\delta}\mathbf{N}\,(r,t)\;E_o\,(r), \quad (3.2)$$

where

$$G\,(R,r)=\frac{e^{ik|\overrightarrow{R-r}|}}{4\pi|\overrightarrow{R-r}|} \quad (3.3)$$

and

$$E_o\,(r)=\sqrt{\frac{P_TG_T}{4\pi}}\frac{e^{i\vec{k}_1\cdot\vec{r}}}{r}. \quad (3.4)$$

P_T is the transmitted power and G_T the transmitter gain function. The integration of (3.2) is extended over the common volume V illuminated by both the transmitter and receiver.

Since both the transmitter and receiver are usually many wavelengths from the scattering blobs in the common volume, one can use the far-field approximation [5] to express the scattered field strength as:

$$E_s(R)\simeq r_0\sqrt{\frac{P_T}{4\pi}}\frac{e^{ikR_2}}{4\pi}\int_V d^3r\,\frac{e^{i\overrightarrow{(k_2-k_1)}\cdot\vec{r}}}{R_1R_2}\boldsymbol{\delta}\mathbf{N}(r,t)\sqrt{G_T} \quad (3.5)$$

where the various quantities are defined in figure 2. The orthogonal random signal component $x(t)$ and $y(t)$ of figure 1 are identified with the real and imaginary parts of this expression respectively. The received power is computed from this expression by averaging its square and multiplying with the effective receiver area. In terms of the scattering difference vector $\vec{K}=\overrightarrow{k_1-k_2}$:

$$P_R=P_Tr_0^2\frac{A_R}{(4\pi)^3}\int d^3r\int d^3r'\,\frac{(G_TG_T'G_RG_R')^{1/2}}{R_1R_1'R_2R_2'}\,e^{(\vec{ir}\cdot\vec{K}-\vec{r}'\cdot\vec{K}')}$$

$$\times<\boldsymbol{\delta}\mathbf{N}(r,t)\boldsymbol{\delta}\mathbf{N}(r',t)>, \quad (3.6)$$

where A_R is the actual receiver area. The space correlation probably vanishes unless r and r' are within several kilometers of one another, and the other factors are relatively weak functions of position, so that $R_1=R_1'$, $R_2=R_2'$, etc., to a good approximation. Transforming the double integral to sum and difference coordinates (r,ρ), one finds:

$$P_R=P_Tr_0^2\frac{A_R}{(4\pi)^3}\int_V d^3r\frac{G_TG_R}{R_1^2R_2^2}\int d^3\rho e^{i\vec{K}\cdot\vec{\rho}}$$

$$<\boldsymbol{\delta}\mathbf{N}(\vec{r},t)\boldsymbol{\delta}\mathbf{N}(\overrightarrow{r+\rho},t)>$$

$$=P_Tr_0^2\frac{A_R}{(4\pi)^3}\int_V d^3r\frac{G_TG_R}{R_1^2R_2^2}\,S(\vec{K}). \quad (3.7)$$

where $S(K)$ is the spectrum of turbulent irregularities to be discussed presently. This expression is probably a valid approximation, at least insofar as single scattering is the dominant scattering mechanism.

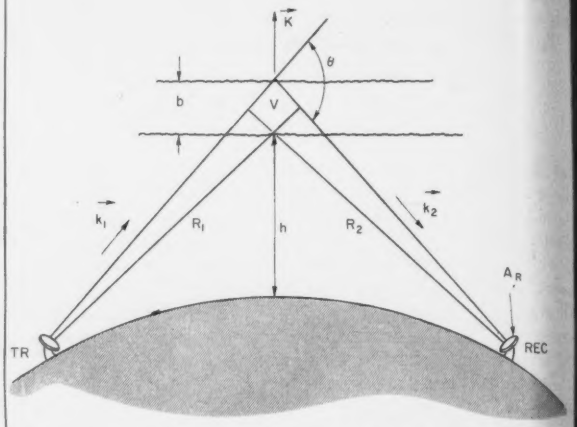


FIGURE 2. *Geometry for analysis of electromagnetic scattering from ionospheric irregularities.*

Transmission expressions which have been used thus far in interpreting experimental data acquired on forward scatter communication circuits, however, represent further assumptions, which are well worth identifying.

a. The ionospheric irregularities are assumed to be isotropic,[3] so that the spectrum $S(K)$ appearing in (3.7) depends only on the magnitude of the scattering difference vector.

$$K=|\vec{K}|=\frac{4\pi}{\lambda}\sin\left(\frac{\theta}{2}\right), \quad (3.8)$$

where θ is the scattering angle between the upgoing and downgoing rays at the integration point r. The scattering process thus acts like a narrow band filter on the spectrum, emphasizing the uniquely important wavenumber given by eq (3.8). The isotropy assumption evidently precludes a discussion of magnetically-alined irregularities with the resulting transmission expression.

b. The volume integration (d^3r) is collapsed by multiplying V by an average value of the integrand. The average scattering angle $\bar{\theta}$ is usually chosen to be that at the path midpoint, corresponding to the strongest scattering contribution (i.e., smallest K). The effective volume is set by $\bar{\theta}$, the transmitter gain, and the width of an effective scattering layer b (see fig. 2):

$$V=\frac{4\pi}{G_T}R_1^2b\,\csc\frac{\bar{\theta}}{2}. \quad (3.9)$$

[3] Anisotropic correlation functions were used by Staras in his extensive analysis of tropospheric scatter circuits and by Booker in discussing the backscattering from field-alined auroral irregularities. However, these techniques have not been applied to ionospheric forward scattering to date.

c. Polarization effects have been neglected, although a relatively simple theory predicts a factor $\sin^2 \chi$ in eq (3.6), where χ is the angle between $\vec{E}_0$ and the scattering direction $\vec{k}_2$. More precise accounts [13, 14] of the vector field response support this approximation.

The resulting transmission equation,

$$P_R = P_T r_0^2 A_R G_R \frac{b}{R_2^2} \csc\left(\frac{\bar{\theta}}{2}\right) S\left(\frac{4\pi}{2} \sin\frac{\bar{\theta}}{2}\right) \quad (3.10)$$

has formed the only linkage between turbulence theory and experiments thus far. Expression (3.10) is evidently a substantial approximation to the actual electromagnetic response of the irregularities, and it is well to be aware of its limitations. On the other hand, a new technique has been developed [5] which allows one to remove all of the above approximations—within the framework of single scattering—at the expense of complicated integrations.

The scattering cross section (per unit volume per unit solid angle) concept is often employed, viz,

$$\sigma(\theta,\lambda) = r_0^2 S\left(\frac{4\pi}{\lambda} \sin\frac{\bar{\theta}}{2}\right). \quad (3.11)$$

However, this approach does not emphasize the *joint* frequency-distance dependence of (3.10) which is characteristic of scatter propagation. In any case, it is quite clear that the spectrum of irregularities is the basic description of the turbulent electron density variations which is required by the electromagnetic theory.

4. Turbulence Theories

Results of the last section show that the spectrum of turbulent electron irregularities is the physical quantity of central interest for scatter propagation. More particularly, one is interested in the spectrum at wavelengths

$$K = \frac{4\pi}{\lambda} \sin\frac{\bar{\theta}}{2}, \quad (4.1)$$

which lie in the following range for typical vhf scatter paths.

$$0.2\text{m}^{-1} < K < 1.3\text{m}^{-1}, \quad (4.1a)$$

On the other hand, the wavenumber K_s at which diffusion-viscosity effects ought to become important is within—or at least near—the range of interest.

$$K_s = \left(\frac{V_0^3}{l_0 \nu^3}\right)^{1/4} \simeq 0.5\text{m}^{-1} \quad (4.2)$$

In terms of the spectral ranges identified in figure 3, this means that one is probably interested in the spectrum for the difficult *transition region* between the inertial subrange and the asymptotic dissipation range. This is unfortunate, since one probably cannot use the dimensional analysis methods which have proven so powerful in the inertial subrange.

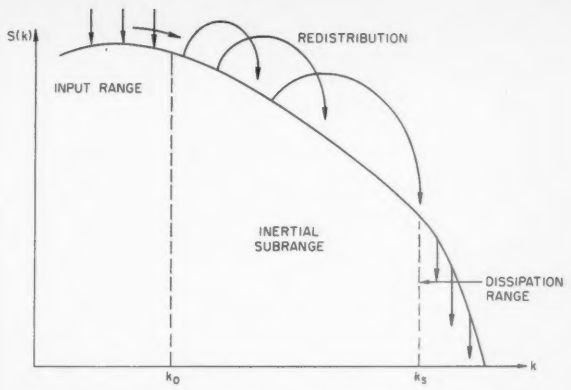


FIGURE 3. *Typical spectrum of dielectric irregularities plotted logarithmically versus wavenumber, identifying various turbulence ranges.*

To avoid notational confusion, we use $S(K)$ to denote the three-dimensional spectrum, which is defined as the Fourier transform of the space correlation function.

$$\langle \delta\mathbf{N}(\bar{r},t)\, \delta\mathbf{N}(\overrightarrow{r+R},t)\rangle = \frac{1}{8\pi^3} \int d^3k e^{\vec{ik}\cdot\vec{R}} S(\vec{K}). \quad (4.3)$$

If the turbulent irregularities are isotropic, the spectrum depends only on the magnitude of the vector wavenumber $\vec{k}$, and the entire description is much simplified.

The one-dimensional spectrum is denoted by $\Gamma(k)$, and corresponds to wavenumber contributions to the mean square variation.

$$\langle \delta N^2 \rangle = \frac{1}{2\pi^2} \int_0^\infty dk k^2 S(\vec{K}) = \int_0^\infty dk \Gamma(k).$$

This establishes the relation:

$$S(\vec{K}) = 2\pi^2 \frac{\Gamma(k)}{k^2}. \quad (4.4)$$

Three distinct physical theories have been developed in recent years to describe the irregularities of electron density established by the turbulent velocity field of neutral ionospheric gases:

a. Theory of pressure fluctuations,
b. Obukhov-Corrsin mixing theory,
c. Villars-Weisskopf mixing theory.

Each of these theories has now been extended into the dissipation regime, and the essential results are tabulated for convenience in figure 4, together with brief references to the many contributors. Without

Theory	Inertial Range	References	Dissipation Range	References
1. Pressure Fluctuations	$N_o\left(\frac{V_o}{C}\right)^4 k_o^{4/3} k^{-13/3}$	(15) Villars and Weisskopf	$N_o\left(\frac{V_o}{C}\right)^4 k_o^{4/3} k_s^{20/3} k^{-11}$	(15) Villars and Weisskopf (16) Wheelon
2. Villars-Weisskopf Mixing	$\left(\frac{dN_o}{dh}\right)^2 k^{-5}$	(17) Villars and Weisskopf (16) Wheelon (19) Wheelon (18) Gallet	$\left(\frac{dN_o}{dk}\right)^2 k_s^8 k^{-13}$	Villars and Weisskopf (Unpublished) (16) Wheelon (19) Wheelon
3. Obukhov-Corrsin Mixing	$\left(\frac{dN_o}{dh}\right)^2 k_o^{-4/3} k^{-11/3}$	(20) Obukhov (21) Corrsin (22) Silverman (23) Bolgiano (24) Batchelor	$\sim k^{-3}$ $D \ll \nu$ $\sim k^{-23/3}$ $D \gg \nu$	(24) Batchelor (25) Batchelor, Howells, and Townsend

FIGURE 4. *Summary of spectral models $S(k)$ for electron density irregularities in the lower ionosphere.*

attempting to plead one case or another in the colorful but unfortunate controversy which has been kindled in this new field, it is important to note several points which now seem rather clearcut:

a. The theory of pressure fluctuations has essentially been abandoned as a major source of ionospheric irregularities [17, 18].

b. The two mixing theories both agree that a (random) convective mixing-diffusion equation of the form:

$$\left(\frac{\partial}{\partial t}-D\nabla^2\right)\delta N(r,t)+\vec{v}(r,t)\cdot\vec{\nabla}\delta N(r,t)=\vec{v}\cdot\vec{\nabla}N_0 \quad (4.5)$$

characterizes the process, where $\vec{v}(r,t)$ is the turbulent solution of the Navier-Stokes equations.

c. The Obukhov-Corrsin theory is patterned precisely on the Kologomoroff theory of the velocity spectrum. It essentially disregards the differential eq (4.5), assumes that the input and inertial range do not overlap, and arrives at the following result by dimensional analysis:

$$\Gamma(k)\sim k^{-5/3}\left(\frac{dN_0}{dh}\right)^2 k_0^{-4/3}. \quad (4.6)$$

d. Villars and Weisskopf noted that the gradient of N_0 appears in the basic eq (4.5). Accepting dN_0/dh as a legitimate parameter for dimensional analysis, one finds:

$$\Gamma(k)\sim k^{-3}\left(\frac{dN_0}{dh}\right)^2. \quad (4.7)$$

e. In deriving expressions for the transition and dissipation ranges, both mixing theories use approximate analysis which are open to question.

f. None of the theories discussed here have as yet been extended to include the influence of the earth's magnetic field.

The reader is referred to the basic references indicated in figure 4 for detailed arguments and precise expressions.

5. Signal Level and Scattering Heights

One can combine the approximate transmission expression (3.10) with various turbulence models listed in figure 4 to predict theoretically the received power on a particular scatter link. This calculation and comparison has been performed for ionosphere forward scatter circuits so far [17, 26] only for the Villars-Weisskopf mixing model (4.7), which is subject to a rather stringent test since it contains no adjustable turbulence parameters. Using only the inertial subrange result for illustration, one has for the theoretical transmission equation:

$$P_R=P_T\frac{bA_RG_R}{R_2^2}2\pi^2l_0^2\left(\frac{\lambda}{4\pi}\right)^5\csc^6\left(\frac{\theta}{2}\right)\left(\frac{dN_0}{dh}\right)^2. \quad (5.1)$$

All of the factors in this equation are known for a given circuit, except the appropriate value of electron density gradient. This evidently depends on the ambient electron density profile at the common volume altitude.

In figure 5 we reproduce a suggested typical noontime profile of electron density for the lower ionosphere from reference [26]. The sharp lower boundary at 70 km is consistent with vlf reflection data, and the upper portions agree moderately well with rocket data. According to both mixing theories, large gradient portions of the profile ought to give rise to more intense turbulent irregularities and hence stronger scattering. A sharp gradient at 70 km would reproduce the strong daytime signal contribution. The power levels measured on the Cedar Rapids-Sterling path [7] require that this

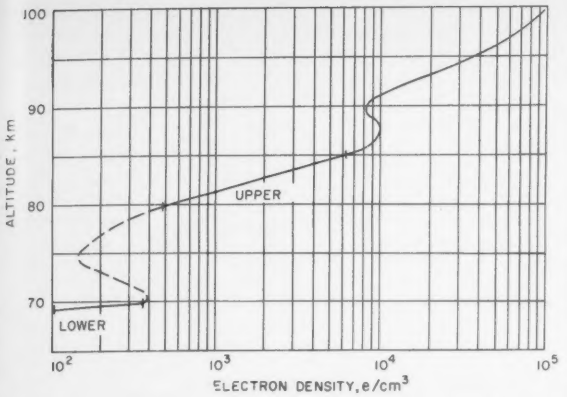


FIGURE 5. *Possible noontime profile of electron density in the lower ionosphere which is consistent with various radio measurements from reference [26].*

gradient be approximately

$$\left.\frac{dN_0}{dh}\right|_{70} \simeq 2{,}000 \frac{e}{\text{cm}^3} \text{ per km}, \qquad (5.2)$$

which represents a very fast rise to the modest total electron density at this height. It would be interesting to see if the Obukhov-Corrsin theory gives a smaller number.

A weaker gradient at approximately 85 km is suggested in figure 5, and should also contribute to the scatter signal. The gradient at this height is thought to be

$$\left.\frac{dN_0}{dh}\right|_{85} \simeq 300 \frac{e}{\text{cm}^3} \text{ per km}, \qquad (5.3)$$

corresponding to a signal some 10- to 15-db weaker than the 70-km contribution. It is interesting that both of these signal components are found in the short pulse measurements of Pineo [27]. In general, he finds unequal signals from two levels separated by about 15 km. Further evidence for weaker scattering from an upper layer comes from the long range NBS experiments [28], which compared 36-Mc signals over 2270 km with 50-Mc signals received on a 1240-km path. Earth screening on the longer path obscured all volume elements below 87 km in these experiments, and the (frequency-scaled) received signal was some 10-db weaker than the shorter circuit, which presumably could exploit both the 70- and 85-km ranges. Further evidence for this dual stratification comes from the vertical incidence pulse backscatter measurements made at hf (see [20] for references).

6. Diurnal and Seasonal Variations

The signal level of a vhf scatter circuit exhibits large changes during each day, year, and sunspot cycle. Experimental data for the NBS 50-Mc Cedar Rapids-Sterling circuit [7] is shown in figure 6, together with the long range Newfoundland-Azores path. The first record is typical of most vhf paths and shows a 10-db rise from morning to noon, followed by a 15-db drop to postsunset values. The corresponding meteor count is strongly peaked at 0600 and is essentially negligible between noon and midnight. This suggests that the afternoon and early evening is under strong solar control.

A theoretical explanation [26] for the afternoon signal drop is offered by turbulent mixing theories through the mean profile gradient factors, since the mean profile is under very strong solar control all day. The noontime signal is contributed by the 70-km level (sec. 5) and the afternoon decay shown in figure 6 ought to reflect the way in which this sharp gradient is erased by recombination at sunset. The rapid drop observed corresponds to the very fast recombination rates at lower levels. On the other hand, the decay of the upper gradient is found to be much more gradual, and in fact should produce measurable scatter signals all night. The noon-to-evening 1500-km signal drop in figure 6 presumably corresponds to the shift of scattering height from the 70- to 85-km levels, which is confirmed by pulse measurements [27]. The comparative absence of a steep signal drop on the long range path is again consistent with a slow recombination of the upper gradient. This line of reasoning forms quite a satisfactory explanation for the existing data on diurnal variations. The basic key to this agreement is the relationship between turbulent intensity and the mean profile provided by the mixing theories.

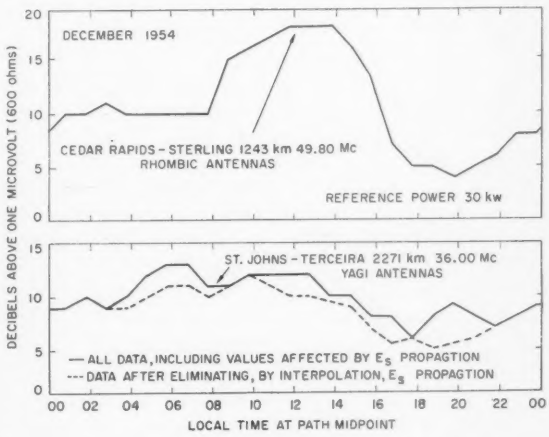

DECEMBER 1954
CEDAR RAPIDS – STERLING 1243 km 49.80 Mc
RHOMBIC ANTENNAS
REFERENCE POWER 30 kw
ST. JOHNS – TERCEIRA 2271 km 36.00 Mc
YAGI ANTENNAS
— ALL DATA, INCLUDING VALUES AFFECTED BY E_s PROPAGTION
--- DATA AFTER ELIMINATING, BY INTERPOLATION, E_s PROPAGTION
DECIBELS ABOVE ONE MICROVOLT (600 ohms)
LOCAL TIME AT PATH MIDPOINT
00 02 04 06 08 10 12 14 16 18 20 22 00

FIGURE 6. *Diurnal variation of power on two vhf scatter paths from reference [7].*

Recent published data [29] on correlation of vhf scatter signal strength and sunspot number demonstrates a strong solar control over a 6-yr period. A good explanation is again provided by the relationship between turbulence intensity and the mean profile (gradient), since the latter is clearly under consistent solar control.

7. Frequency Dependence

The basic transmission eq (3.10) depends upon the radiofrequency through: (1) The turbulence spectrum $S(K)$, and (2) the effective receiving area $A_R G_R$. To avoid confusion on how the second effect is accounted for, it is convenient to compare received power on two different frequencies for *scaled* antennas, viz, equal G_R factors. The definitive experiments by the NBS group [10] on frequency dependence actually used carefully scaled rhombic antennas, so as to insure identical illumination of the scattering layers and similar antenna response. With this convention, the frequency variation is that due to the spectrum plus a λ^2 factor which represents the physical scaling of the receiving aerial size. If the spectrum is expressible as a simple inverse power (n) of wavenumber in the relevant turbulent range, one would predict a (scaled antenna) frequency scaling law [30],

$$P_R = \frac{\text{const}}{f^{n+2}}, \tag{7.1}$$

The most recent experimental data is that reported by Blair [10], using 30, 40, 50, 74, and 108 Mc. He finds that the exponent n in eq (7.1) is variable, ranging from 7 to 10. This variation is both diurnal and seasonal, so that it is dangerous to quote a mean value. During late afternoon hours, however, when scattering from turbulence is presumably the dominant mode, the exponent n is definitely higher, and a value of 8 to 9 may be representative. Comparing this with the turbulence results tabulated in figure 4, the Obukhov-Corrsin inertial range result gives $f^{-5.6}$, which is definitely too low. The Villars-Weisskopf inertial result f^{-7} is occasionally correct, but the corresponding dissipation behavior f^{-15} is too high. The $f^{-9.6}$ dissipation result for $d >> \nu$ given by Batchelor, Howells, and Townsend [25] describes some of the records. On the balance, however, one finds that none of the theories give a satisfactory explanation for the important frequency exponent variability.

Earlier experiments of frequency dependence [7, 31] indicated a higher exponent scaling law between 50 and 100Mc, than that measured between 30 and 50Mc. Because the scattering process acts like a narrow band filter on the spectrum (sec. 3), it was suggested that the viscosity transition wavenumber of the ionosphere (4.2) did, in fact, divide the wavenumber range of experimental interest (4.1). The steeper slope approaching the dissipation range and the smaller inertial slope seemed to explain the data rather nicely [16]. However, the newer data [10] indicates that such a dichotomy does not exist, and that the exponent is constant over the wavenumber range employed at any given hour—even though the slope itself shows diurnal and seasonal variations.

8. Distance Dependence

If the turbulent spectrum function $S(K)$ is expressible as a simple inverse power (n) of wavenumber in the relevant range, two factors which depend on path length d occur in the basic transmission eq (3.10),

$$P_R = \frac{\text{constant}}{R_2^2 \left(\sin\frac{\theta}{2}\right)^{n+1}} \tag{8.1}$$

If a denotes the mean radius of the earth, the distance R_2 from the average scattering point to the receiver is given by

$$R_2^2 = a^2 + (a+h)^2 - 2a(a+h)\cos\left(\frac{d}{2a}\right) \simeq \left(\frac{d}{2}\right)^2 \text{for } a \gg d. \tag{8.2}$$

The scattering angle is a more complicated function of distance d and scattering height h [16].

$$\tan\frac{\theta}{2} = \frac{1-\cos\left(\frac{d}{2a}\right)+\frac{h}{a}}{\sin\left(\frac{d}{2a}\right)} \simeq \frac{d}{4a} + \frac{2h}{d} \text{ for } d \ll a. \tag{8.3}$$

A graphical evaluation of Θ versus d for various h is given in reference [16]. By combining all of the above expressions, it is possible to establish the variation of received power with distance predicted by any turbulence model (n). However, one should also note that earth screening horizon limitations have the effect of raising the lowest scattering midpoint as the distance is increased. Insofar as the strongest scattering centers are located at the lower height (sec. 6), this vertical inhomogeneity must be reflected in an additional distance dependence of the mean signal level.

The experimental data now available is probably not accurate enough for a decisive test of the theories. The early data [7] used successive measurements of 50-Mc signals at 491, 592, and 811 km, whereas one would have preferred simultaneous data over longer paths. The shipborne DSIR experiments [31] on 41 Mc gave limited results in terms of defining the distance scaling law [30]. Recent (unpublished) airborne receiver experiments made by the Lincoln Laboratory group out to 3000 km on 50 Mc should provide a valuable basis for testing the various theoretical results.

9. Sudden Ionospheric Disturbances

A sudden increase in the flux of ionizing radiation from the sun increases the electron density rather quickly in the lower ionosphere, which in turn increases the nondeviative absorption of radio waves which pass through the D region [33]. This prompt absorption is called a sudden ionospheric disturbance, and affects both totally reflected short wave (hf) signals and the partially reflected waves responsible for (vhf) ionospheric scatter propagation. A typical set of simultaneous signal amplitude records made on a shortwave circuit at 6 Mc and two adjacent scatter links at 28 and 50 Mc is reproduced in figure 7, from reference [7]. The shortwave link drops out immediately, while the 50-Mc scatter link increases

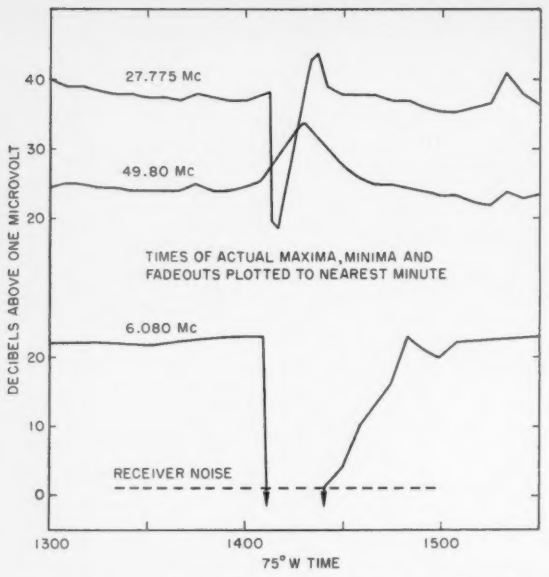


FIGURE 7. *Transient behavior of power on three frequencies taken from reference* [7].

gradually until the recovery cycle of the transient D region begins. The 28-Mc link shows a mixture of the two effects. Viewed in a qualitative way, these data lend strong support to the notion that ionosphereic scatter links are under direct solar control much of the time [29].

One can construct [4] an analytical explanation for the transient amplitude behavior shown in figure 7. Nondeviative absorption of a radio wave of frequency f can be represented approximately by a multiplicative factor [33];

$$\exp -\frac{\gamma}{f^2}\int_0^h dy N(y,t)\nu(y), \qquad (9.1)$$

where h is the reflection height. $N(y,t)$ is the instantaneous value of electron density at the ray path integration point y, and ν is the collision frequency at the same point. The totally reflected shortwave field strength is modified by the factor (9.1) *alone*, which is very large because the carrier frequency f is relatively small. The rapid realization of this absorption loss indicated by figure 7 indicates how rapidly the electron density profile is modified by the burst of solar ionization.

If one is to believe the turbulent mixing theories, the field strength of an ionospheric scatter circuit depends both on the transient electron density profile and its gradient, viz,

$$P \simeq \left[\frac{dN(h,t)}{dh}\right]^2 \exp -\frac{\gamma}{f^2}\int_0^h dy N(y,t)\nu(y). \qquad (9.2)$$

[4] A. D. Wheelon, unpublished calculations.

Because of height-dependent recombination rates, etc., rapid changes in the flux of ionizing radiation produce substantial modification of the basic electron density height profile. Careful analysis shows that transient modification of ionization gradients in the D region can explain the signal increase of 50 Mc shown in figure 7. (The exponential factor is negligible at such high frequencies.) On the other hand, the 28-Mc signal ought to suffer significant absorption, and the competition between enhanced gradient scattering and this absorption can be explained with the theory. A *possible* explanation for this type of behavior is thus provided by turbulent mixing theories and standard absorption expressions.

10. References

[1] K. Bullington, Characteristics of beyond-the-horizon radio transmission, review paper, Proc. IRE **43,** 1175 (1955).

[2] H. Bremmer, Terrestrial radio waves (Elsevier, New York, N.Y., 1949).

[3] H. G. Booker and W. E. Gordon, A theory of radio scattering in the troposphere, Proc. IRE **38,** 401 (1950)

[4] C. L. Pekeris, Note on the scattering of radiation in an inhomogeneous medium, Phys. Rev. **71,** 268 (1947).

[5] A. D. Wheelon, Radiowave scattering by tropospheric irregularities, J. Research NBS **63D,** 205 (1959).

[6] D. K. Bailey, R. Bateman, L. V. Berkner, H. G. Booker G. F. Montgomery, E. M. Purcell, W. W. Salisbury, and J. B. Wiesner, A new kind of radio propagation at very high frequencies observable over long distances, Phys. Rev. **86,** 141 (1952).

[7] D. K. Bailey, R. Bateman, and R. C. Kirby, Radio transmission at vhf by scattering and other processes in the lower ionosphere, Proc. IRE **43,** 1181 (1955).

[8] W. G. Abel, J. T. deBettencourt, J. H. Chisholm, and J. F. Roche, Investigations of scattering and multipath properties of ionospheric propagation at radio frequencies exceeding the MUF, Proc. IRE **43,** 1255 (1955).

[9] R. A. Silverman, Some remarks on scattering from eddies, Proc. IRE **43,** 1253 (1955).

[10] J. C. Blair, Frequency dependence of vhf ionospheric scattering, unpublished work (1959).

[11] H. Staras, Diversity reception with correlated signals, J. Appl. Phys. **27,** 93 (1956). See also Proc. IRE **44,** 1057 (1956), and Study and investigation of tropospheric scattering, RCA Rept. Pt. A, p. 91 (1956).

[12] A. D. Wheelon, Refractive corrections to scatter propagation, J. Geophys. Research **62,** 343 (1957).

[13] M. Balser, Some observations on scattering by turbulent irregularities, Trans. IRE PGAP **AP-5,** 383 (1957).

[14] S. Stein, Some observations on scattering by turbulent irregularities, Trans. IRE PGAP **AP-6,** 299 (1958).

[15] F. Villars and V. F. Weisskopf, The scattering of electromagnetic waves by turbulent atmospheric fluctuations, Phys. Rev. **94,** 232 (1954).

[16] A. D. Wheelon, Radio frequency and scattering angle dependence of ionospheric scatter propagation at vhf, J. Geophys. Research **62,** 93 (1957).

[17] F. Villars and V. F. Weisskopf, On the scattering of radio waves by turbulent fluctuations of the atmosphere, Proc. IRE **43,** 1232 (1955).

[18] R. M. Gallet, Aerodynamical mechanisms producing electron density fluctuations in turbulent ionized layers, Proc. IRE **43,** 1240 (1955). See also, The spectrum of the electron density fluctuations in the ionosphere, p. 165 in Polar Atmos. Symp., Pt. 2, Ionospheric Sec. (Pergammon Press Inc., New York, N.Y.)

[19] A. D. Wheelon, Spectrum of turbulent fluctuations produced by convective mixing of gradients, Phys. Rev. **105,** 1706 (1957).
[20] A. M. Obukhov, Structure of the temperature field in turbulent flow, Izvest. Akad. Nauk SSSR, Ser. Geog. Geofiz. **13,** 58 (1949).
[21] S. Corrsin, On the spectrum of isotropic temperature fluctuations in an isotropic turbulence, J. Appl. Phys. **22,** 469 (1951).
[22] R. A. Silverman, Turbulent mixing theory applied to radio scattering, J. Appl. Phys. **27,** 699 (1956).
[23] R. Bolgiano, The role of turbulent mixing in scatter propagation, Trans. IRE PGAP **AP–6,** 159 (1958).
[24] G. K. Batchelor, Small scale variation of convected quantities like temperature in turbulent fluid, I, J. Fluid Mech. **5,** 113 (1959).
[25] G. K. Batchelor, I. D. Howells, and A. A. Townsend, Small scale variations of convected quantities like temperature in turbulent fluid, Pt. II, J. Fluid Mech. **5,** 134 (1959).
[26] A. D. Wheelon, Diurnal variations of signal level and scattering heights for vhf propagation, J. Geophys. Research **62,** 255 (1957).
[27] V. Pineo, Oblique incidence measurements of the heights at which ionospheric scattering of vhf radio waves occur, J. Geophys. Research **61,** 165 (1956).
[28] R. C. Kirby, Extreme useful range of vhf transmission by scattering from the lower ionosphere, IRE Conv. Record, Pt. I, p. 112 (1958).
[29] C. Ellyett and H. Leighton, Solar cycle influence on the lower ionosphere and on vhf forward scatter, Proc IRE **46,** 1711 (1958).
[30] H. Ekre, K. Endresen, T. Hagfors, B. Landmark, and J. Rodsrud, A study of ionospheric vhf forward scattering at high latitudes, Norweg. Defense Research Establ. Rept. No. 27, Oslo, Norway.
[31] W. J. Bray, J. A. Saxton, R. W. White, and Luscombe, vhf propagation by ionospheric scattering and its application to long-distance communication, Proc. IEE, 1033, p. 236 (Oct. 1955).
[32] S. K. Mitra, The Upper Atmosphere, Calcutta, Asiatic. Soc. Bengal, 2d ed. Asiatic Soc. Monograph Ser. **5,** (1952).

(Paper 64D4–62)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Propagation at Oblique Incidence Over Cylindrical Obstacles[1]

M. P. Bachynski

(November 24, 1959; revised January 11, 1960)

Investigations of propagation of short electromagnetic waves at oblique incidence over smooth, perfectly conducting cylindrical obstacles are described. It is shown that the effect of oblique incidence can be considered as a change in the effective radius of curvature of the diffracting obstacle. The power in the shadow region of a cylindrical obstacle decreases with angle of obliqueness for horizontally polarized waves and can decrease, remain constant, or increase with angle of obliqueness for vertically polarized waves depending on the geometry of the propagation link. In all cases, vertical polarization gives a stronger field in the shadow region than horizontal polarization. In addition it is shown that the diffracted field behind an obstruction of uniform radius of curvature is the same as that behind an obstacle of uniformly varying radius of curvature, provided the effective radius is the same.

1. Introduction

Of common occurrence in radio engineering, is the situation where an obstruction lies across the direct line path between transmitting and receiving stations. If the obstacle is smooth, of uniform radius of curvature, and is located normal to the line joining the transmitter and receiver, then its effect has been investigated by several authors [1, 2, 3, 4] [2] who studied diffraction by smooth cylindrical obstacles at grazing angles. In particular the latter two [3, 4] brought their investigations to the point to be useful for numerical calculations. However, if the obstacle is located at an oblique angle to the direct line path and/or is better approximated by a cylinder of varying radius of curvature along its length, then the application of the above analysis is not straightforward. In fact, few investigations [5, 6] have been extended to include the case when the electromagnetic energy is other than normally incident on the obstacle.

One case of field measurements has been reported [7] where a diffracting ridge was inclined under an angle of 60° to the path of radiation. However, at that time, no theory was available to take the obliqueness of the obstacle into account. Furthermore, no experimental evidence of the effect of oblique incidence appears to have been published to date.

In this paper, propagation of electromagnetic energy at oblique incidence over smooth, perfectly conducting cylindrical obstacles of both uniform and linearly varying radius of curvature is investigated. The experimental results are compared to a generalized theory of diffraction by smooth conical obstacles [6].

[1] Contribution from Research Laboratories, RCA Victor Company, Ltd., Montreal, Canada.

[2] Figures in brackets indicate the literature references at the end of this paper.

2. Theory

In figure 1, T and R represent transmitter and receiver locations respectively. Let a cylindrical obstacle of uniform height horizontally, but varying

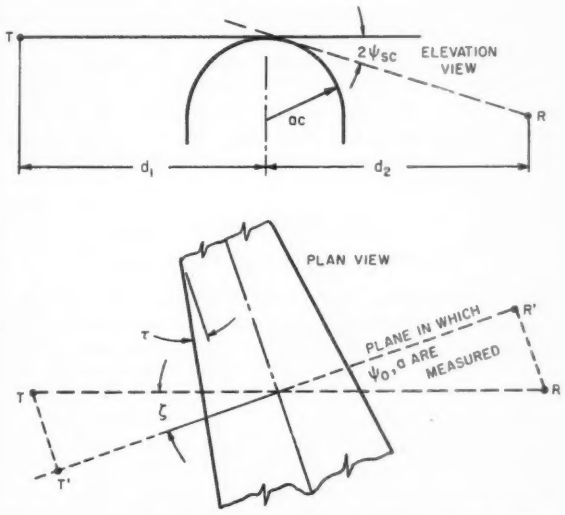


FIGURE 1. *Notation for propagation at oblique incidence over a cylindrical obstacle.*

radius of curvature (cone angle 2τ), cross the propagation path so that the normal to the axis of the cylindrical obstacle is inclined at an angle ζ (angle of obliqueness) to the straight line joining the transmitter-receiver. The scattering angle and radius of curvature in the vertical plane containing the re-

ceiver and transmitter are then:

$$2\ \psi_{sc}=2\ \cos\ \zeta\ \cos\ \tau\ \psi_0$$

$$a_e=a/(\cos\ \zeta\ \cos\ \tau)$$

where $2\psi_0$, a are the scattering angle and radius of curvature, respectively, measured in the plane perpendicular to the axis of the cylinder and passing through the point where a vertical plane containing the transmitter and receiver would intersect the crest of the obstacle. (In general, the transmitter-receiver does not lie in this plane [6] which defines ψ_0 and a.)

The results for diffraction by cylindrical obstacles [3] at normal incidence have been shown to be directly applicable [6] for conical obstacles at oblique incidence provided the following substitutions are made:

$$2\psi_0 \rightarrow 2\psi_{sc}=2\ \cos\ \zeta\ \cos\ \tau\ \psi_0 \qquad (1)$$

$$a \rightarrow a_{eff}=a/\cos^2\ \zeta\ \cos\ \tau \qquad (2)$$

Hence it now becomes convenient, for the case of oblique incidence and/or varying radius of curvature to speak of an effective radius of curvature a_{eff} given by eq (2). The scattering angle ψ_{sc} is the actual scattering angle determined by the geometry of the propagation path.

Equation (2) indicates that the effect of oblique angles of incidence are of much greater influence than a changing radius of curvature of the intervening obstacle. Thus the distinction between a cylindrical obstacle of constant radius of curvature and one of linearly varying radius of curvature (half-cone) will not be significant except for obstructions of large cone angles ($>30°$) which are seldom found in practice. As can be seen the variations become more pronounced as the angle of obliqueness increases. These predictions will be verified in the following section.

3. Experiment

3.1. Apparatus and Technique

Measurements of power diffracted by smooth, perfectly conducting cylinders of both uniform and linearly varying radius of curvature were conducted within the laboratory using model techniques described previously [3, 6]. Uniform cylinders with ka values of 0 (knife edge), 154, 202 and 254 and cylinders of varying radius of curvature (half-cone) with 15° half-cone angle, and ka values of 130, 194 and 254 were used in the experiments.

Measurements of diffracted power as function of angle of obliqueness could be performed by rotating the model about a vertical axis in the vertical plane containing the fixed receiver-transmitter positions and passing through the axis of the model. In this manner, ka_{eff} values greater than 2,000 could be obtained. The sizes of the models permitted investigations for angles of obliqueness of up to 60° from the normal before end effects became important. In addition, power measurements behind obstacles for scattering angle of up to 12° were made.

In all measurements, the top surface of the diffracting cylinder was horizontal, and the incident electric vector orientated so as to be either parallel (horizontal polarization) or perpendicular (vertical polarization) to the grazing surface. In this way, cross-polarization effects were minimized. As well, since the obstacles were perfectly conducting, no cross-polarized effects due to oblique incidence of the radiation should be present [5].

3.2. Power Variation With Angle of Obliqueness

The power variation at different receiver heights with angle of obliqueness (ζ) is shown in figure 2

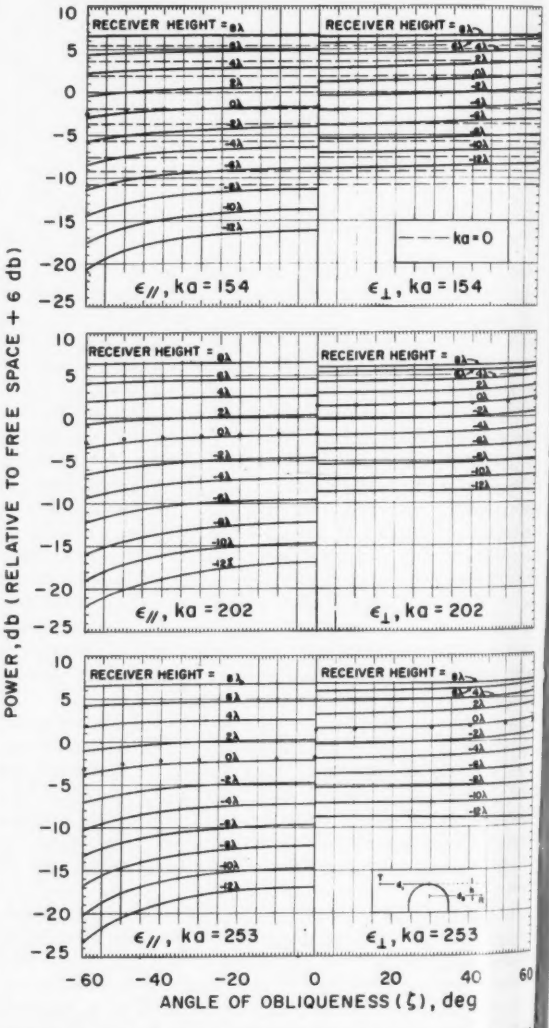

RECEIVER HEIGHT =
ka = 0
$\epsilon_{//}$, ka = 154
$\epsilon_{\perp}$, ka = 154
$\epsilon_{//}$, ka = 202
$\epsilon_{\perp}$, ka = 202
$\epsilon_{//}$, ka = 253
$\epsilon_{\perp}$, ka = 253
POWER, db (RELATIVE TO FREE SPACE + 6 db)
ANGLE OF OBLIQUENESS (ζ), deg

FIGURE 2. *Experimental measurements of power variation with angle of obliqueness for a smooth, perfectly conducting cylindrical obstacle of constant radius of curvature for both horizontal polarization* ($\epsilon\parallel$) *and vertical polarization* ($\epsilon\perp$).

The dashed lines are for a knife edge: $ka \rightarrow 0$. The dots represent theoretical values. ($d_1=150\lambda$, $d_2=113\lambda$, $\lambda=1.25$ cm.)

(cylindrical obstacles of uniform radius of curvature) and figure 3 (obstacles of linearly varying radius of curvature) for obstacles of various size.

a. Vertical Polarization

For vertical polarization, the diffracted power in the receiving space varies with angle of obliqueness as follows:

1. Illuminated region ($2\psi_{sc}<0$, corresponding to receiver heights $h>0$). The diffracted power increases with angle of obliqueness and increasing ka_{eff}. Small values of ka_{eff} show little variation with angle of obliqueness.

2. Halo region ($2\psi_{sc}\sim 0$–$5°$, $h\sim 0$ to -5λ). Power increased with increasing angle of obliqueness and increasing ka_{eff}.

3. Shadow region ($2\psi_{sc}>5°$, $h<-5\lambda$). The power deep in the shadow depends on the value of ka_{eff}. For small ka_{eff} values, the power continues to increase with increasing angles of obliqueness while for large values of ka_{eff} the diffracted power decreases with angle of obliqueness. Alternately, regions are possible where the diffracted power is independent of angle of obliqueness.

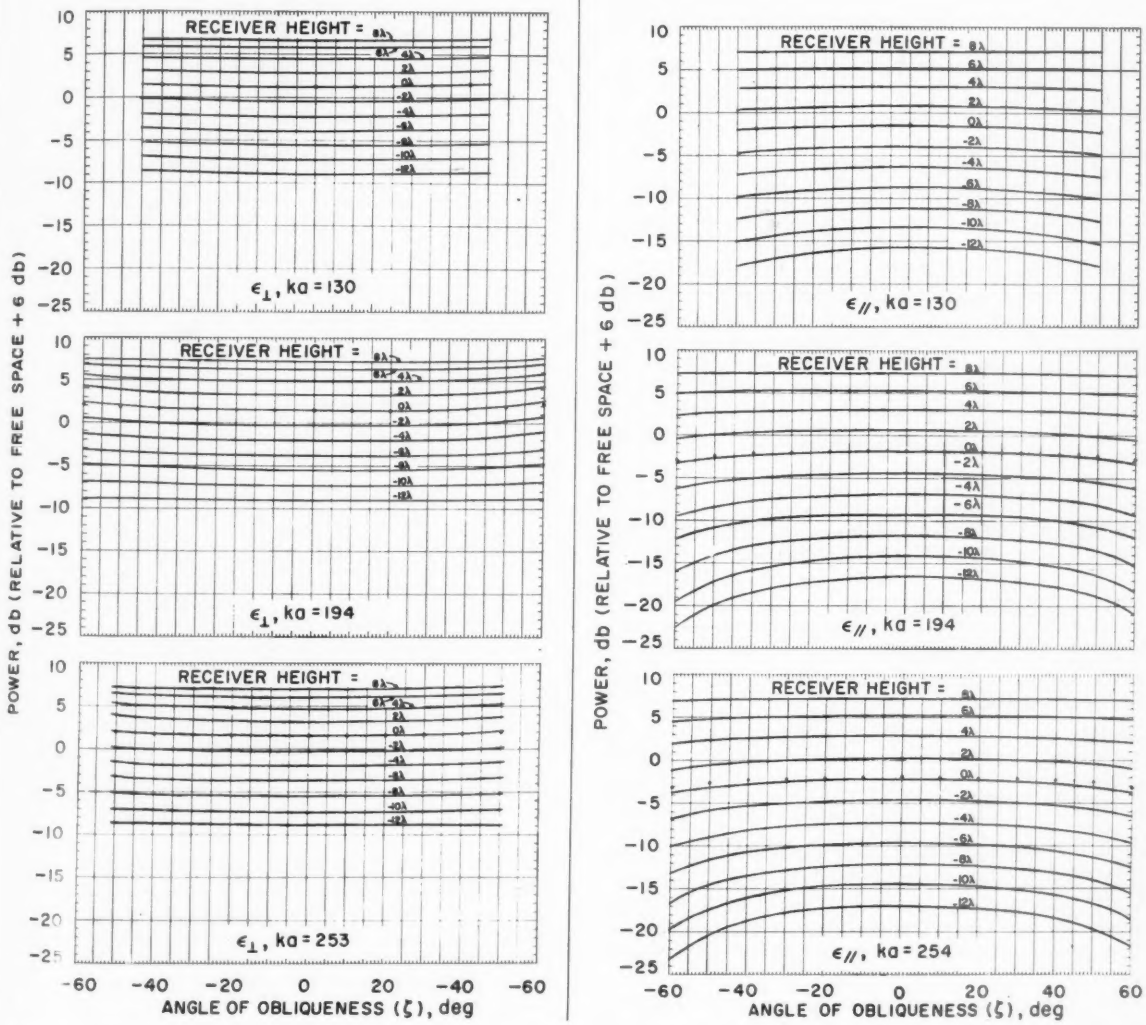

RECEIVER HEIGHT =
$\epsilon_\perp$, ka = 130
$\epsilon_\perp$, ka = 194
$\epsilon_\perp$, ka = 253
$\epsilon_{/\!/}$, ka = 130
$\epsilon_{/\!/}$, ka = 194
$\epsilon_{/\!/}$, ka = 254
POWER, db (RELATIVE TO FREE SPACE + 6 db)
ANGLE OF OBLIQUENESS (ζ), deg

FIGURE 3. *Experimental measurements of power variation with angle of obliqueness for a smooth perfectly conducting obstacle of uniformly varying radius of curvature (15° half cone angle) for both horizontal polarization ($\epsilon_{/\!/}$) and vertical polarization ($\epsilon_\perp$).*

The dots represent theoretical values. ($d_1=150\lambda$, $d_2=113\lambda$, $\lambda=1.25$ cm.)

b. Horizontal Polarization

For horizontally polarized waves incident on the obstacles the diffracted power behavior is:

1. Illuminated region—the power is nearly constant with angle of obliqueness, showing only a small increase at large ka_{eff} values and large oblique angles.
2. Halo region—the diffracted power decreases with increasing angle of obliqueness and increasing ka_{eff} values.
3. Shadow region—the power decreases rapidly with angle of obliqueness and increasing effective radius.

The diffracted power behind a knife edge obstruction ($ka_{eff}=0$) is also shown in figure 2. It should be noted that the diffracted power in this case is independent of both the angle of obliqueness of the incident energy and of polarization.

Comparison of the data for obstacles of constant radius and varying radius of curvature shows that the diffracted power variation with angle of obliqueness is indeed similar provided their effective radii of curvature (a_{eff}) are the same. In all cases the horizontally polarized field results in the diffracted power in the halo and shadow regions being considerably less than for a knife edge obstruction. The vertically polarized energy, on the other hand, is greater than the values for a knife edge obstacle in the halo region but can be less than the knife edge values deep in the shadow region. These results will be discussed further in section 3.3. Theoretical values of the diffracted power for grazing incidence ($\psi_{sc}=0$) are shown in each case and are found to agree well with experiment up to angles of obliqueness exceeding 50°.

3.3. Effect of Oblique Incidence on Power Variation With Scattering Angle

The variation of the diffracted power with scattering angle is shown for normal incidence ($\zeta=0°$) and an angle of obliqueness of 60° for obstacles of constant radius in figure 4 and for obstacles of varying radius of curvature in figure 5.

In the illuminated region, the horizontally polarized energy is considerably greater than the vertically polarized energy. The power for a knife edge obstruction is intermediate to the values for the two polarizations. The effect of oblique incidence is small for either polarization in the illuminated region.

For small scattering angles, the diffracted power for vertical polarization is greater at oblique angles, while the reverse is true for horizontal polarization. Comparison with theory shows satisfactory agreement.

At large scattering angles, the vertically polarized energy decreases with angle of obliqueness and becomes *less* than for normal incidence. The horizontally polarized energy, however, continues to decrease more rapidly with angle of obliqueness. This behavior for large *ka* values has been predicted for

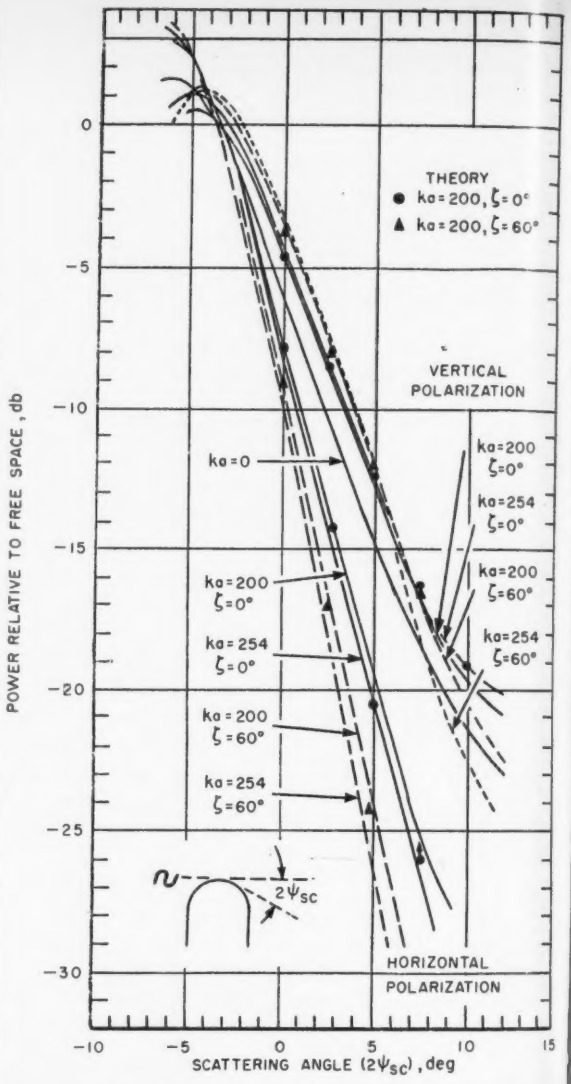


FIGURE 4. *Variation of diffracted power with scattering angle for angles of incidence of 0 and 60°.*

Obstacles of constant radius of curvature. ($d_1=150\lambda, d_2=113\lambda, \lambda=1.25$ cm.)

normal incidence in earlier publications [2, 8] but not verified experimentally. The theory for diffraction by conical obstacles [6] appears to hold up to scattering angles of 5 deg or more, then deviates from experimental measurements as the geometric approximations used in numerical evaluation of the theory become significant.

The power distribution behind a uniform cylindrical obstacle has been found experimentally to be the same as the power distribution behind an equivalent cylinder of uniformly varying radius of curvature.

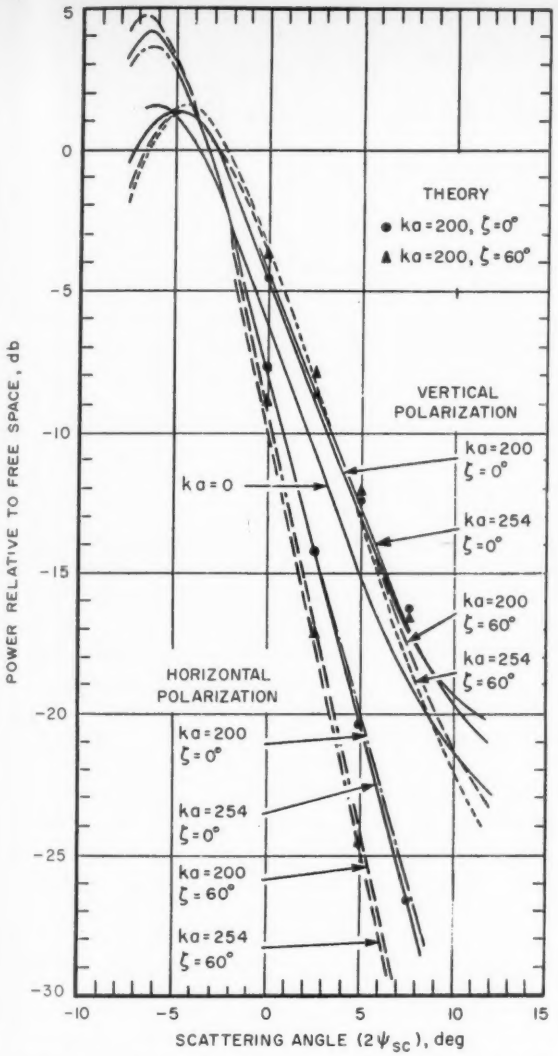


FIGURE 5. *Variation of diffracted power with scattering angle for angles of incidence of 0 and 60°.*

Obstacles of uniformaly varying radius of curvature −15° half-cone angle. (d_1=150λ, d_2=113λ, λ=1.25 cm.)

(Paper 64D4–63)

4. Conclusions

It is shown that, for propagation of electromagnetic waves at oblique incidence over smooth, perfectly conducting cylindrical obstacles, the effect of oblique incidence can be accounted for by assigning an effective radius of curvature to the obstacle and then considering propagation to occur at normal incidence.

The effect of oblique incidence is to sharply decrease the scattered power for horizontally polarized energy, while for vertically polarized energy, the diffracted power is enhanced for small scattering angles and decreased at large scattering angles—a behavior similar to that predicted [2, 8] for normal incidence. In all cases the power in the halo and shadow region is less for horizontal polarization than for vertical polarization.

There is no difference in behavior of the power whether it is diffracted by an obstruction of uniform radius of curvature or by an obstruction which has a varying radius of curvature (half-cone) providing the ka_{ett} values are the same. For small grazing angles the theory for diffraction by smooth conical obstacles [6] agrees well with experimental measurements.

The author is indebted to the Air Force Cambridge Research Center for support under Contract No. AF.19(604)–3049, and to M. G. Kingsmill for performing the experimental measurements. Discussions with Dr. H. E. J. Neugebauer have provided much insight into this work.

5. References

[1] V. Fock, The distribution of currents induced by a plane wave on the surface of a conductor, J. Phys. (U.S.S.R.) **10,** 130 (1946); The field of a plane wave near the surface of a conducting body, J. Phys. (U.S.S.R.) **10,** 399 (1946).

[2] S. O. Rice, Diffraction of plane radio waves by a parabolic cylinder, Bell System Tech. J. **33,** 417 (1954).

[3] H. E. J. Neugebauer and M. P. Bachynski, Diffraction by smooth cylindrical mountains, Proc. IRE **46,** 1619 (1958).

[4] J. R. Wait and A. M. Conda, Diffraction of electromagnetic waves by smooth obstacles for grazing angles, J. Research NBS **63D,** 181 (1959).

[5] J. R. Wait, Scattering of a plane wave from a circular dielectric cylinder at oblique incidence, Can. J. Phys. **33,** 189 (1955).

[6] H. E. J. Neugebauer and M. P. Bachynski, Diffraction by smooth conical obstacles, J. Research NBS **64D,** 316 (1960).

[7] J. H. Crysdale, et. al., An experimental investigation of the diffraction of electromagnetic waves by a dominating ridge, Trans. IRE AP–**5,** 203 (1957).

[8] I. P. Shkarofsky, M. P. Bachynski, and H. E. J. Neugebauer, Electromagnetic wave propagation over mountains, Can. IRE Conv. Record (Oct. 1958).

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Diffraction by Smooth Conical Obstacles [1]

H. E. J. Neugebauer [2] and M. P. Bachynski

(November 24, 1959; revised January 11, 1960)

Expressions obtained earlier [1, 2] [3] for the calculation of diffraction due to conducting obstacles with smooth cylindrical surfaces, are generalized to oblique incidence and to surfaces of conical shape. The derivation is based on a generalized concept of the Green's function and on the use of corrective factors that take the same place as corrections introduced by other authors into the theory of diffraction by apertures. The final expressions for conical obstacles and oblique incidence are very similar to those for cylindrical obstacles. The results are compared with scale model measurements.

1. Introduction

Kirchhoff's theory for the solution of diffraction problems can be generalized to be applicable to the case of scattering by cylindrical obstacles with smooth surfaces. In two recent publications [1, 2] it has been shown that the predictions of this theory are in good agreement with experiments. In addition, J. R. Wait and A. M. Conda [3] have derived numerical results from a different theory which are also in agreement with our own experimental and theoretical conclusions.

An object of the present paper is a further generalization of our investigations to conical obstacles and oblique incidence. Another object is a more thorough justification of the theoretical procedure used previously [1].

2. Basic Theory

In figure 1, T_1 represents a transmitter, T_2 a receiver, and Z a perfectly conducting obstacle of cylindrical shape between the two stations. Green's theorem (1) can be used to calculate the field E at receiver T_2.

$$E=-\frac{1}{4\pi}\iint\left(G\frac{\partial E'}{\partial n}-E'\frac{\partial G}{\partial n}\right)dS. \tag{1}$$

The integration is extended over the entire plane Ξ separating the two stations, the normal n points inwards, E' is the radiation at the plane Ξ and G is Green's function. If the field E' and Green's function were rigorously known the field E could be rigorously calculated from (1). Neither E' nor G are known, but approximate values can be obtained by the following reasoning.

Incident radiation travels directly from T_1 to V. In addition, some radiation is reflected at point S_1 of the obstacle and reaches point V. The sum of these two radiations is used as E' in eq (1).

Such a procedure obviously involves two approximations: (a) Strictly speaking the radiation reflected from a curved surface cannot be calculated from geometrical optics. However, the curvature of the obstacle is assumed to be small enough so that geometrical optics is a permissible approximation. (b) As is typical for Kirchhoff's theory, the unperturbed incident field is substituted for the true field in the plane of integration.

---

[1] Contribution from Research Laboratories, RCA Victor Company, Ltd., Montreal, Canada.
[2] Present address: 59 Fuller Ave., P.O. Box 176, Webster, N.Y.
[3] Figures in brackets indicate the literature references at the end of this paper.

Turning next to the discussion of Green's function, we recall that G is the field at T_2 when a unit source of radiation is located at V. When the effect of obstacles is neglected

$$G=\frac{e^{-jkR_2}}{R_2} \tag{2}$$

where R_2 is the distance from V to T_2, and $k=2\pi/\lambda$ is the wave number. This basic form of Green's function is used for the solution of many simple diffraction problems.

The correct form of Green's function must, however, satisfy the boundary conditions at the obstacle Z [4 to 7]. In the case of plane obstacles the image of V is added as a virtual source of radiation. In the present case, an approximate form of Green's function is obtained by adding to the expression of eq (2), the reflected radiation traveling from V via S_2 to T_2. A necessary precondition for such a procedure is that geometrical optics is a sufficiently accurate approximation for the calculation of the reflected radiation.

Another justification for the use of this approximation follows from the property of Green's function to be symmetrical in the coordinates of V and T_2. Placing a unit source of radiation at T_2 and calculating its field at V is essentially the same problem as calculating the incident field at V due to a source at T_1. Hence, it is reasonable to use the same method for the determination of E' and G, thereby automatically satisfying the reciprocity theorem.

Summarizing the preceding considerations we state that

$$E'=\frac{e^{-jkR_1}}{R_1}+\rho\ \mathrm{Div}_1\frac{e^{-jk[R_1'+r_1]}}{R_1'+r_1} \tag{3}$$

$$G=\frac{e^{-jkR_2}}{R_2}+\rho\ \mathrm{Div}_2\frac{e^{-jk[R_2'+r_2]}}{R_2'+r_2} \tag{4}$$

can be expected to be useful approximations for the field in plane Ξ and for Green's function.

The meaning of the distances R_1, R_1', r_1, . . . can be taken from figure 1. The symbols Div_1 and Div_2 denote divergence factors [8] that take care of the loss of intensity when radiation is reflected by a curved surface. The factor ρ is different for vertical and horizontal polarization. It will be further discussed at the end of this section and in section 3.2(d).

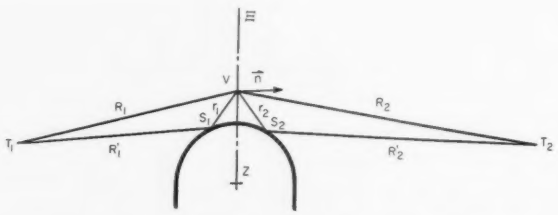


FIGURE 1. *Notation for diffraction by a cylindrical obstacle.*

In applying equation (1) use will be made of the inequality

$$kr\gg 1 \tag{5}$$

so that the derivatives of E' and G are

$$\frac{\partial E'}{\partial n}=-jk\frac{e^{-jkR_1}}{R_1}\cos(R_1,n)-jk\,\rho\ \mathrm{Div}_1\frac{e^{-jk[R_1'+r_1]}}{R_1'+r_1}\cos(r_1,n) \tag{6}$$

$$\frac{\partial G}{\partial n}=-jk\frac{e^{-jkR_2}}{R_2}\cos(R_2,n)-jk\,\rho\ \mathrm{Div}_2\frac{e^{-jk[R_2'+r_2]}}{R_2'+r_2}\cos(r_2,n). \tag{7}$$

Substitution of (3), (4), (6), and (7) into (1), yields the sum of four terms, the first one being exactly the same as when the obstacle is a conducting half plane. This term is called knife edge term. It is characterized by the path of radiation $T_1V\,T_2$.

The other terms which are called halo terms, can be characterized in a similar manner by the paths of radiation which are $T_1S_1V\,T_2$, $T_1V\,S_2T_2$, $T_1S_1V\,S_2T_2$. It is seen that the treatment of the problem is very similar to what is generally known as four-ray-theory [9] of a knife edge obstacle over plane conducting ground. In that theory the factors ρ have different values for parallel and perpendicular polarization. For a perfectly conducting plane earth,

$$\rho_{\parallel}=-1 \tag{8}$$

$$\rho_{\perp}=+1 \tag{9}$$

Since geometrical optics has been used to calculate the field reflected by the obstacle it would appear that the same values must be used for ρ in connection with the present theory. However, a better approximation can be obtained based on the succeeding arguments.

As mentioned at the beginning of this section, a correct value for E would be obtained from (1) if the correct field values of E' were known in plane Ξ. Instead the values of the incident field are substituted for E'. In the analogous case of diffraction by a circular aperture, Braunbek [10] could improve the results by adding, to the incident field, field values near the rim of the aperture which he derived from Sommerfeld's [11] electromagnetic theory of the diffraction field near the straight edge of an infinite half-plane. The same approach has been used by H. Levine [12] in his treatment of the same problem.

In the present case of scattering by a cylindrical obstacle a similar improvement can be obtained when electromagnetic theory of scattering by a conducting cylinder is combined with the observation that the main contributions to the integrals representing the halo terms are supplied by rays that are almost grazing. [1].

From Fock's [13] investigations it is known that, in the case of perpendicular polarization, the electric field at the grazing point is very nearly 1.4 times the incident field, whereas straight geometrical optics would yield a factor of two in agreement with (9). Hence, it is to be expected that

$$\rho_{\perp}=0.7 \tag{10}$$

leads to a better agreement with experimental evidence than $\rho_{\perp}=1.0$.

This reasoning is unsatisfactory only in one respect. It cannot be applied to the case of parallel polarization because the total field at the surface of a perfect conductor is zero whether it is calculated by geometrical optics or by electromagnetic theory. Hence, any finite value could be chosen for $\rho_{\parallel}$. The only justification for setting

$$\rho_{\parallel}=-1.0$$

is that it is the simplest assumption and that it leads to good agreement with experimental results. However, a slightly smaller value than 1.0 might produce an even better agreement.

Strictly speaking ρ is not a constant but depends on the location of the point of the surface where the reflection takes place. Thus, the incident field yielding the knife edge term would also have to be multiplied by a similar factor ρ' since the strength of this field is also reduced very near the top of the obstacle where, for perpendicular polarization, $\rho'=\rho=0.7$. Since the zone of the plane of integration where $\rho'\neq 1$ is narrow compared with the first Fresnel zone the factor ρ' can be set equal to unity for the calculation of the knife edge term. On the other hand, the first "Fresnel zone" in the calculation of the halo terms is so small that ρ can be set equal to the constant value 0.7. It has been shown [14] that the agreement between theory and experiment can be slightly improved by calculating the halo terms with a varying ρ factor.

3. Detailed Theory

The general theory of the previous section may now be applied to generalize the earlier results [1, 2] in two directions: (a) The mountain may have a top that is more appropriately described by a cone. (b) Electromagnetic radiation is incident under an angle that may be different from 90 degs.

An approximate description of the obstacle would be that of a cone with a fairly small cone angle, with its axis more or less horizontal, and with its circular base approximately vertical. The tip of the cone is, in general, far away from the point where the radiation crosses the obstacle (fig. 2).

The theory will be general enough to cover also another case of great practical interest (fig. 3). The base of the cone is essentially horizontal, its axis approximately vertical, and the radiation is diffracted around the obstacle at two opposite sides. In this case the cone angle is, in general, fairly large.

It may be mentioned that the picture of "creeping waves" is not commensurate with this theory. As is typical for Huygen's Principle, all the rays from the source reaching the plane of integration, either directly or after reflection, generate secondary radiation. The sum of all these secondary wavelets at the receiver is the field E of eq (1). Hence, the question for the path of radiation along the surface of the scatterer is not meaningful in this theory.

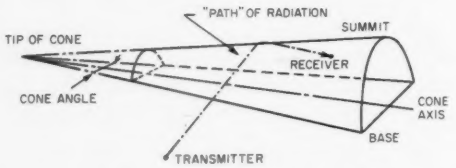


FIGURE 2. *Diffraction by a conical obstacle where radiation travels around one side of the cone.*

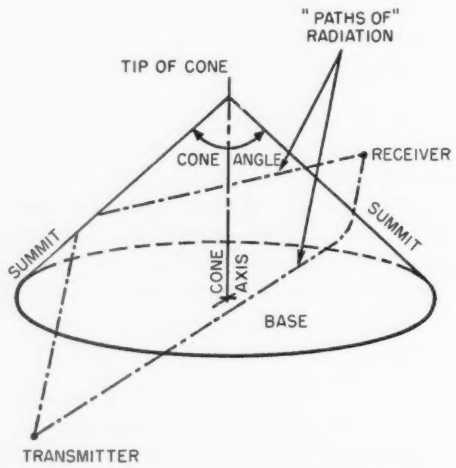


FIGURE 3. *Diffraction by a conical obstacle where radiation can travel around both sides of the cone.*

3.1. Knife Edge Term

The knife edge term corresponds to the "direct" path of radiation ($T_1V\,T_2$) and is obtained by replacing the obstacle by a conducting half-plane. For the situation of figure 2, the resulting obstacle is a simple, rectangular knife edge for which the solution is well known. The knife edge obstacle for the case of figure 3, represents a plane obstacle of triangular shape. The diffraction of electromagnetic waves by such an obstacle will be dealt with in the sequel.

Figure 4 represents a plane conducting obstacle of triangular shape as seen from the transmitter T. The straight line connecting transmitter T and receiver T_2, intersects the obstacle at B. The obstacle is subdivided into five different areas whose effects are calculated separately and by different methods.

Lines b and b' are drawn, through B, parallel to the edges a and a' of the obstacle. They include the angle ϕ. Lines BC and BC' are perpendicular to a and a' respectively. Finally

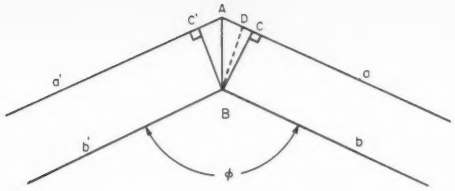


FIGURE 4. *Subdivision of triangular region into five different areas.*

A and B are connected by a straight line. Temporarily it is assumed that TT_2 is perpendicular to the plane of the obstacle. This simplifying assumption will be dropped later.

If there is no obstacle the total unobstructed field radiated from T towards T_2 is well known. The field intercepted by the obstacle can be calculated by means of Kirchhoff's formula and subtracted from the total field (Babinet's Principle). For this purpose the obstacle is subdivided into: (1) The triangular area, of infinite extent, bounded by b and b'; (2) the two rectangular areas bounded by a, b, and BC and by a', b', and BC'; and (3) the two triangles ABC and ABC'.

The contribution of the triangular area bounded by b and b' is $\phi/2\pi$ of the total field. This follows immediately from Kirchhoff's formula. The total field without obstacle is given by

$$E_{\text{total}}=\frac{jk}{4\pi}\int_0^{2\pi}\int_0^{\infty}\frac{e^{-jk(R_1+R_2)}}{R_1R_2}(\cos\theta_1-\cos\theta_2)\,\rho\,d\rho\,d\phi \tag{11}$$

where: the integration is extended over the entire plane, ρ, ϕ are cylindrical coordinates with origin at B. As usual, R_1 and R_2 are the distances from T to the element of integration and from there to T_2. θ_1 and θ_2 are the angles of incidence. When a triangular obstacle bounded by b and b' is present, the upper limit of the first integral of eq (11) is $(2\pi-\phi)$ instead of 2π. Since the integrand is independent of ϕ, eq (12) is obtained.

$$E=E_{\text{total}}-\frac{\phi}{2\pi}E_{\text{total}}. \tag{12}$$

The effect of the rectangular areas can be calculated by exactly the same method used in the well known treatment of Fresnel diffraction by a half plane. For normal incidence it is found that the effect of the rectangular area bounded by the straight lines a, b, and BC, is given by

$$E_{\text{rect.}}=e^{-jkd_m}\cdot\left(\frac{1+j}{4d_m}\right)\left[\left(C\sqrt{\frac{2k}{\pi d_m}}\,c\right)-jS\left(\sqrt{\frac{2k}{\pi d_m}}\,c\right)\right] \tag{13}$$

where C and S are Fresnel's cosine and sine integrals, the meaning of c is explained by figure 5, and that of d_m by eq (18).

To obtain the "knife edge" term of the scattered field it is necessary to subtract, from the right-hand side of eq (12), the expression (13), with a corresponding one for the other rectangular area and expressions for the effect of the two finite triangular areas ABC and ABC'. Triangle ABC may be treated as example. The angles of incidence can be considered constant when the dimensions c and s of the triangle (fig. 5) are small compared to the distance d_1 and d_2 of the stations from point B. For perpendicular incidence, $\theta_1=\theta_2-\pi=0$. The (negative) contribution of the triangle is

$$E_{\Delta}=\frac{jk}{2\pi d_1d_2}\iint e^{-jk(R_1+R_2)}\,d\sigma \tag{14}$$

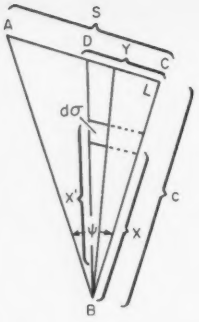


FIGURE 5. *Notation for evaluating contribution of area* A B C.

where,

$$d\sigma=\frac{x}{c}\,dy\,dx. \tag{15}$$

Making the substitution:

$$R_1+R_2=d+\frac{x^2}{d_m}\,(1+y^2/c^2) \tag{16}$$

with

$$d=d_1+d_2; \tag{17}$$

$$2/d_m=1/d_1+1/d_2 \tag{18}$$

and integrating with respect to x, we obtain:

$$E_\Delta=\frac{e^{-jkd}}{2\pi dc}\int_0^s \frac{dy}{1+y^2/c^2}\left[1-e^{-jk\frac{c^2}{d_m}(1+y^2/c^2)}\right]. \tag{19}$$

The integral splits up into the difference of two integrals. The second one is evaluated by the stationary phase method yielding

$$E_\Delta=\frac{e^{-jkd}}{2\pi dc}\left[c\psi-\sqrt{\pi d_m/2k}\,e^{-\left(jk\frac{c^2}{d_m}\right)}\left\{C(\sqrt{2k/\pi d_m}\,s)-jS(\sqrt{2k/\pi d_m}\,s)\right\}\right]. \tag{20}$$

Minor changes take place when the plane of the obstacle is not perpendicular to the straight connection TT_2 For the succeeding derivations it is assumed that the decisive contribution to the integrals in question are made by surface elements $d\sigma$ whose distance ρ from B is small compared to d_1 and d_2 so that the angle of incidence ζ can be considered constant for the integration, ζ being the angle between TT_2 and the perpendicular of the plane of the obstacle. With this assumption, the integrand of eq (11) can be written

$$2\,\frac{e^{-jk(R_1+R_2)}}{R_1R_2}\cos\zeta d\sigma. \tag{21}$$

The product $\cos\zeta\ d\sigma$ can be replaced by $d\sigma'$ where $d\sigma'$ is the surface element obtained by projection of $d\sigma$ onto a plane perpendicular to TT_2.

Figure 6 represents a plane through the stations T, T_2, and through the element of integration $d\sigma$. The distance of $d\sigma$ from B is ρ. Depending on the location of $d\sigma$ in the plane of the obstacle, angle Ω assumes a value between $\pi/2-\zeta$ and $\pi/2+\zeta$. An elementary calculation yields

$$R_1+R_2=d_1+d_2+(\rho^2/d_m)\sin^2\Omega.$$

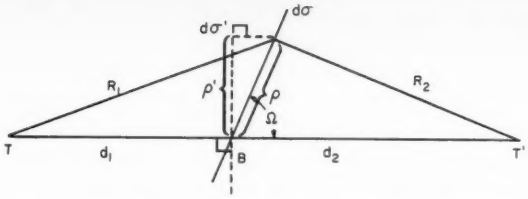


FIGURE 6. *Plane through station* TT′ *for calculation of effect of oblique incidence on triangular obstacle.*

Denoting the distance of $d\sigma'$ from B by $\rho' = \rho \sin \Omega$ we can replace (21) by

$$2 \frac{e^{-jkd}}{d_1 d_2} e^{-jk\rho'^2/d_m} d\sigma'. \tag{22}$$

This expression establishes the rule that a plane oblique obstacle must be projected onto a perpendicular plane and that the diffraction field must be calculated as if the projection were the obstacle.

3.2. Halo Terms

The halo terms represent the additional radiation due to the smooth surface of the obstacles rather than the sharp edges of their plane substitutes. The following derivations will be based more on the situation of figure 2, rather than figure 3, although the theory applies to both cases.

If a_c is the radius of curvature, it will be assumed that $ka_c >> 1$. This would exclude the area near the tip of the cone which may be important in the case of figure 3. The halo terms obtained for the cylindrical mountain in the earlier publications [1, 2]tend towards zero with vanishing a_c in agreement with experiment. Since this theory will lead to a similar result we feel certain that it can be applied to any part of the cone, be it near its tip or far away from it, although $ka_c >> 1$ is a basis of its derivation.

Experimental evidence does not indicate the existence of a special contribution of the tip of the cone of any appreciable amount. This is in agreement with the general experience that a point singularity does not contribute appreciably to the far field.

It is understandable that the mathematical expression will be more involved for oblique incidence on a conical mountain than for normal incidence on a cylindrical mountain, the latter being a special case of the former. The more important difference, however, is that the integral over the mathematical plane which is typical of Huygen's principle, reduces immediately to a simple integral in the cylindrical case whereas the double integral must be treated much more carefully in the conical case.

The reason is that the situation of the point of stationary phase for the integration parallel to the cylinder axis is self evident. For the conical obstacle and oblique incidence, the point of stationary phase must be determined by a rather laborious calculation. This will be described in several steps.

a. Geometrical Preliminaries

Following notations will be used: Point 0 is the tip of the cone; cone angle is 2τ; and transmitter and receiver are located at points T_1 and T_2, respectively.

Two tangential planes, one through T_1 and another one through T_2, touch the cone in the generatrices OU_1 and OU_2 where the exact location of points U_1 and U_2 along the generatrices will be determined later. The two tangential planes T_1OU_1 and T_2OU_2 intersect in a straight line ON. The plane through the cone axis and through ON will be used as plane Ξ of integration in the sense of Huygen's principle. A third tangential plane is perpendicular to Ξ and is used as base plane.

Figure 7 shows a ground plan of stations and conical obstacle. This figure is obtained by projection onto the base plane. The projections of points onto the base plane are marked by brackets added to the notation of the corresponding points. The straight connection of points (T_1) and (T_2) intersects the plane of integration in A so that OA is the generatrix that lies in the base plane. The acute angle between $(T_1)A$ and the perpendicular to OA is called "angle of incidence ζ." The lengths of $\overline{T_1(T_1)}$ and $\overline{T_2(T_2)}$ are denoted h_1 and h_2 (fig. 8). They are taken positive for points under the base plane.

A plane through A and perpendicular to the cone axis intersects the cone in a circle whose center is M and whose radius is a. Points U_1, U_2, and N are located in this plane.

It is assumed that the distance $d_1=\overline{(T_1)A}$ and $d_2=\overline{(T_2)A}$ are large compared to a,

$$d_1>>a; \qquad d_2>>a. \tag{23}$$

A right-handed rectangular coordinate system with origin in 0 is used. The positive z-axis is OA, the x-axis lies in the base plane, the y-axis in the plane of integration, $OMAN$.

Figure 8 shows an intersection with the plane MU_1AU_2 which is perpendicular to the cone axis. We set

$$\angle AMU_1=\angle AMU_2=\psi_0 \tag{24}$$

Another point on the circumference of the circle is denoted S_1 and,

$$\angle U_1MS_1=\psi. \tag{25}$$

A ray transmitted from T_1 and incident at S_1 is reflected by the cone so that it intersects the plane of integration, $x=O$, in V. Ray S_1V seems to come from the image T_1' of point T_1. Points T_1 and T_1' are symmetrical to the tangential plane through generatrix OS_1. (In general, none of the points T_1, T_1', V lies in the plane of fig. 8.)

Any other ray transmitted from T_1 and incident at one of the points of generatrix OS_1, is reflected as if it comes from T_1' and intersects the plane of integration in a point of the straight line OV.

Generatrix OA will be called summit of the mountain. The radius of curvature of the mountain at point A is obtained by putting a plane perpendicular to the summit through A, which intersects the cone axis in the center of the osculating circle M'. Hence, the radius of curvature in this cross section is $z_0 \tan \tau=a/\cos \tau$. The radius of curvature in the cross section of a plane through T_1AT_2 (which contains also M') is

$$a_e=a/(\cos \tau \cos \zeta). \tag{26}$$

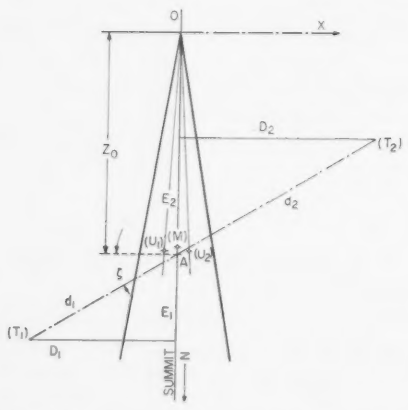


FIGURE 7. *Plan view of stations and conical obstacle.*

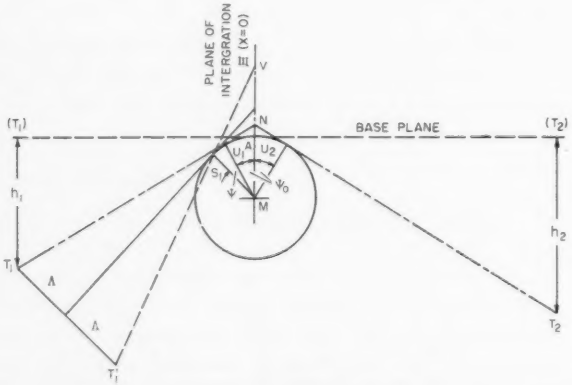


FIGURE 8. *View in plane perpendicular to cone axis.*

The angles of incidence occurring in eqs (6) and (7) can be approximated by ζ or $\pi+\zeta$, respectively. Substituting (3), (4), (6), and (7) into (1), we obtain

$$E=\frac{jk}{4\pi}\{e_0+\rho e_1+\rho e_2+\rho^2 e_{12}\} \tag{27}$$

where e_0, e_1, e_2, e_{12} are integrals characterized by the paths of radiation T_1VT_2, $T_1S_1VT_2$, $T_1VS_2T_2$, $T_1S_1VS_2T_2$. Being slowly varying functions the distances R_1, $R_1'+r_1$, R_2, $R_2'+r_2$ in the denominators of (3), (4), (6), and (7) can be approximated by d_1 and d_2, respectively.
Hence,

$$e_0=\frac{2\cos\zeta}{d_1d_2}\iint e^{-jk[R_1+R_2]}\,dS \tag{28}$$

$$e_1=\frac{2\cos\zeta}{d_1d_2}\iint e^{-jk[R_1'+r_1+R_2]}\,dS \tag{28'}$$

$$e_{12}=\frac{2\cos\zeta}{d_1d_2}\iint e^{-jk[R_1'+r_1+R_2'+r_2]}\,dS. \tag{28''}$$

The expression for e_2 is obtained from (28′) by an exchange of subscripts "1" and "2".

b. Determination of the point of stationary phase

In this subsection a series of calculations will be briefly described which are required for the evaluation of the integrals (28′) and (28″). The integral for e_1 will be considered in particular.

Point V has been obtained by reflection of radiation at point S_1, another point V' is obtained by reflection at a slightly different point whose corresponding angle is $\psi+d\psi$ instead of ψ.

The double integral for the summation of the contribution of all the wavelets is carried out in two steps, first a simple integral over a narrow triangular strip VOV' which extends to infinity beyond V and V', followed by another simple integral over the different strips from $\psi=0$ up.

The integration over one strip is done by the stationary phase method. The point of stationary phase is obtained by putting a plane through OV and T_2 and folding it up by tilting it about OV until it coincides with plane OVT_1'. By this operation point T_2 is brought to T_2^*. The straight line $T_1'\,T_2^*$ intersects OV at the point of stationary phase and the length of $\overline{T_1'\,T_2^*}$ determines the phase of the radiation arriving from T_1 at T_2 after reflection from the cone and reradiation.

The details of the calculations shall not given here.[4] They are simplified by the assumption of small scattering angles so that powers of higher than first order in ψ_0 are neglected. From the corresponding calculations [1] for normal incidence on cylindrical mountains it is concluded that powers up to the third in ψ must be carried. Mixed terms in $\psi_0\,\psi^3$ have been disregarded as the numerical calculations for the cylinder [1] have shown that their effect on the result is negligible.

c. Evaluation of the Integrals

The substitution (29) can be made in the exponent of eq (28′),

$$R_1'+r_1+R_2=\text{Dist }(T_1'T_2^*)+\delta\ \text{D.st }(T_1'T_2^*). \tag{29}$$

It can be shown that [4]

$$R_1'+r_1+R_2=\text{Dist }(T_1'T_2^*)+\frac{\cos^2\zeta}{d_m}\,\delta\xi^2. \tag{30}$$

[4] For details, see: RCA Victor Research Rept. No. 7-100, 4.

When corresponding substitutions are made in (28) and (28'') the eqs (31), (32), and (33) are obtained.

$$e_0=2\frac{\cos\zeta}{d_1d_2}\iint\exp\left[-jk\left\{\text{Dist }(T_1VT_2)+\frac{\cos^2\zeta}{d_m}\delta\xi^2\right\}\right]dS \tag{31}$$

$$e_1=2\frac{\cos\zeta}{d_1d_2}\iint\text{Div }(S_1)\exp\left[-jk\left\{\text{Dist }(T_1'T_2^*)+\frac{\cos^2\zeta}{d_m}\delta\xi^2\right\}\right]dS \tag{32}$$

$$e_{12}=2\frac{\cos\zeta}{d_1d_2}\iint\text{Div }(S_1)\text{ Div }(S_2)\exp\left[-jk\left\{\text{Dist }(T_1'T_2^*)+\frac{\cos^2\zeta}{d_m}\delta\xi^2\right\}\right]dS. \tag{33}$$

The integration of the factor of (32) that depends on ξ' where:

$$\xi'=\sqrt{\frac{2k}{\pi d_m}}\cos\zeta\,d\xi$$

is achieved by the stationary phase method.

$$\int_{-\infty}^{\infty}\sqrt{\frac{\pi d_m}{2k}}\frac{1}{\cos^2\zeta}\exp\left(-\frac{jk}{d_m}\cos^2\zeta\,\delta\xi^2\right)d\xi'=\sqrt{\frac{\pi d_m}{k}}\frac{\exp(-j\pi/4)}{\cos^2\zeta}.$$

Substituting this expression into (32) together with some added manipulations,

$$\frac{jk}{4\pi}e_1=\sqrt{\frac{kd_m}{\pi}}\frac{a_e\exp(j\pi/4)}{2d_1d_2\cos\zeta}\int\sqrt{\chi(3\chi+2\psi_{sc})}\exp[-jk\text{ Dist }(T_1'T_2^*)]d\chi \tag{34}$$

where:

$$\text{Dist }(T_1'T_2^*)=d+2(a_e/\cos\zeta)\chi^3+5(a_e/\cos\zeta)\psi_{sc}\chi^2, \tag{35}$$

and

$$\chi=(\cos\zeta\cos\tau)\psi.$$

Comparison of (34) and (35) with eqs (13b) and (14) through (17) of [1] yields immediately the result that the first halo term is exactly the same in both cases if a is replaced by $a_e/\cos\zeta$ and ψ_0 by ψ_{sc}

The third halo term represented by e_{12} is calculated in a corresponding manner. As has been pointed out [1] the angles ψ on the two sides of the plane of integration are not, in general, equal in magnitude. However, a mean value for ψ is assumed to be sufficiently accurate.

It can be found that

$$\text{Dist }(T_1'T_2'^*)=d+4(a_e/\cos\zeta)\chi^3+7(a_e/\cos\zeta)\psi_{sc}\chi^2. \tag{36}$$

Again a comparison with eqs (21) and (22) of [1] shows that the third halo term is the same if a is replaced by $a_e/\cos\zeta$ and ψ_0 by ψ_{sc}.

A more accurate calculation yields an additional term to the expression (36) for Dist $(T_1'T_2'^*)$,

$$[3\,a_e^2/(d_m\cos^2\zeta)]\,\psi_{sc}\chi^3.$$

d. Correction Factors

A last remark shall be made with respect to the correction factors $\rho_\parallel$ and $\rho_\perp$ which were discussed at the end of section 2. It has still to be shown that $\rho_\perp=+0.7$ also for oblique incidence.

A convenient way makes use of formulas published by J. R. Wait [15, 16]. Wait studies oblique incidence of a plane wave on a cylinder by means of a series expansion. Since he considers the case of parallel polarization, the easiest way of using his equations is to consider the perpendicular component of the magnetic vector H_ρ^s for a cylinder of infinite permeability which behaves like the electric vector for infinite conductivity. With Wait's notations, from Maxwell's equations,

$$-kj\,\eta\,H_\rho^s=\partial E_z^s/\partial(\rho\phi)-\partial E_\phi^s/\partial Z.$$

Substituting from his eqs (3) and (6) for E_z^s and E_ϕ^s and setting $\phi=\pi/2$, $\rho=a$,

$$H_\rho^s=-\frac{E_0}{\eta v}\sum\frac{J_n'(v)}{H_n^{(2)'}(v)}H_n^{(2)}(v)\exp(-jkz\cos\theta+j\phi t)$$

where J_n and $H_n^{(2)}$ are the Bessel and Hankel functions in the usual notations, and his v is, with our notations,

$$v=ka\sin\zeta$$

It follows that,

$$|H_\rho^s|=-\frac{E_0}{\eta v}\left|\sum\frac{J_n'(v)}{H_n^{(2)'}(v)}H_n^{(2)}(v)\right|.$$

Fock [13] has investigated the case $\zeta=\pi/2$ and $ka>>1$. He found a numerical value for $|H_\rho^s|/E_0$ which is independent of a. Hence, $|H_\rho^s|/E_0$ is independent of v and, consequently, it is also independent of ζ when $\zeta\neq\pi/2$.

4. Experiment

4.1. Apparatus and Technique

Experimental measurements of power diffracted by smooth half-conical mountains have been performed in the *K*-band frequency range using model techniques. The details of the experimental arrangement have been discussed previously [1].

Two overlapping, perfectly conducting, conical mountains of 15° half-angle were used in the investigations. The first was connected to a knife edge obstruction with a transition into a 15° cone which extended to a ka of about 200. The second was a complete half-cone and covered the range of $ka=0$ to 400. The knife edge cone transition was used in order to provide an automatic normalization of the received power as a function of effective radius of the cone to the power diffracted by a knife-edge ($ka=0$) obstacle. Thus a continuous plot of diffracted power for different ka values ranging from 0 to 330 was possible by moving the cones perpendicular to the plane joining a fixed transmitter-receiver location before end effects became of consequence. In all measurements the surface of the cone (and the knife edge) was either parallel or perpendicular to the incident electric vector so that cross-polarization effects would not be present.

4.2. Diffraction by Conical Mountains

The power behind a conducting conical mountain as function of ka_e($k=2\pi/\lambda$, a_e= effective radius of the cone in the vertical plane containing both transmitter and receiver) for normally incident ($\zeta=0$) electromagnetic waves at the grazing angle is shown in figure 9. A large difference between the two polarizations exists with the vertically polarized field which increases with radius of curvature being much stronger than the horizontally polarized field which decreases with radius of curvature (ka_e). The measurements agree well with the expression for power at grazing incidence derived earlier.

A complete family of measurements of power variation with ka_e when the electromagnetic energy is normally incident on a conical mountain for different receiver positions (i.e., different scattering angles) is shown in figure 10a, for vertical polarization, and in figure 10b for horizontal polarization.

The variation of power diffracted by a conical obstacle with scattering angle is shown in figure 11. As for a cylindrical mountain, the received power for vertical polarization increases

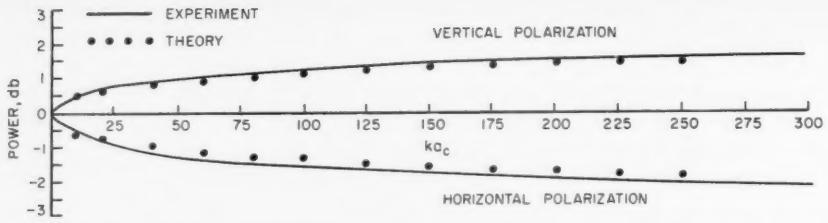


FIGURE 9. *Variation of power behind a conical mountain at grazing incidence with radius of curvature of the obstacle in the vertical plane containing the transmitting and receiving terminals.* (d_1=150λ, d_2=113λ, ψ_o=0°, λ=1.252 cm.)

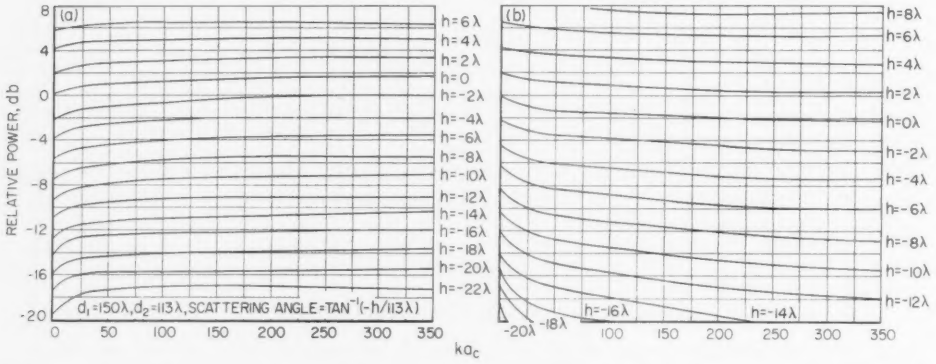


FIGURE 10. *Measured variation of power behind a conical mountain with radius of curvature of the obstacle in the vertical plane containing the transmitting and receiving terminals.*

a, vertical polarization and b, horizontal polarization.

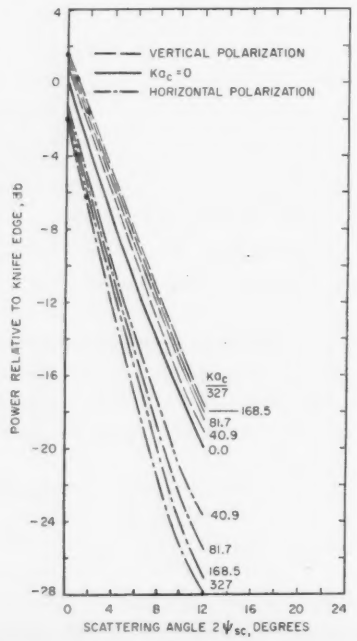


FIGURE 11. *Variation of power behind a conical mountain with scattering angle as a function of radius of curvature.* (d_1=150λ, d_2=113λ, $\psi_{sc}=\psi_0 \cos\tau \cos\zeta$; $a_c=a/\cos\tau\cos\zeta$; τ=15°, ζ=0°.)

Theoretical points are for ka_c=239.

with ka_e while the reverse is true for horizontal polarization. In addition, the slope of the power variation with scattering (diffraction) angle becomes steeper with increasing curvature for horizontal polarization and remains essentially the same as for a knife edge in the case of vertical polarization. The agreement with theory appears satisfactory.

The power distribution behind a conical mountain for $ka_e=327$ is shown in figure 12. These should be compared to the power variation behind a knife edge obstacle [1]. Calculations of the power for grazing incidence are shown and agree well with the measurements.

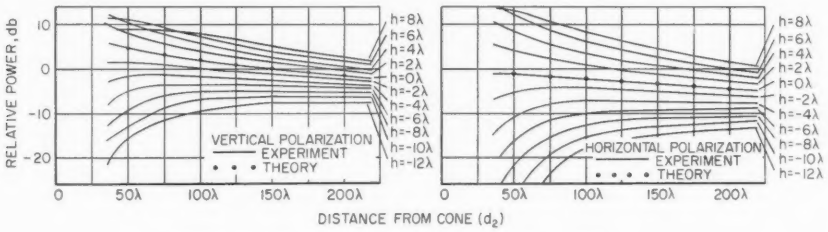


FIGURE 12. *Power distribution behind a conical mountain.*
(k a_e=327, d_1=150λ, λ=1.252 cm.)

The authors thank the Air Force Cambridge Research Center for supporting this work under Contract AF19(604)3049, and M. G. Kingsmill, for performing the experimental measurements.

5. References

[1] H. E. J. Neugebauer and M. P. Bachynski, Diffraction by smooth cylindrical mountains, Proc. IRE **46,** 1619 (1958).
[2] I. P. Shkarofsky, H. E. J. Neugebauer, and M. P. Bachynski, Effect of mountains with smooth crests on wave propagation, IRE Trans., PGAP **AP–6,** 341 (1958).
[3] J. R. Wait and A. M. Conda, Diffraction of electromagnetic waves by smooth obstacles for grazing angles, J. Research NBS **63D**, 181 (1959).
[4] A. Sommerfeld, in Frank-Mises, Die Differential—und Integralgleichungen der Mechanik und Physik, 2d ed. **II,** p. 853 (Braunschweig, 1935).
[5] P. M. Morse and H. Feshbach, Methods of theoretical physics, **I,** p. 812 (McGraw-Hill Book Co., New York, N.Y., 1953).
[6] H. Severin, Zur Theorie der Beugung elektromagnetischer Wellen, Z. Phys. **129,** 426 (1951).
[7] W. Franz, Ueber die Greenschen Funktionen des Zylinders und der Kugel, Z. Naturforsch. **9a,** 705 (1954).
[8] H. J. Riblet and C. B. Baker, A general divergence formula, J. Appl. Phys. **19,** 63 (1948).
[9] J. C. Schelling, C. R. Burrows, and E. B. Ferrell, Ultra-short wave propagation, Proc. IRE **21,** 427 (1933).
[10] W. Braunbek, Neue Näherungs Methode für die Beugung am ebenen Schirm; Zur Beugung an der Kreisscheibe; Z. Physik **127,** 381–390 and 405–415 (1950).
[11] M. Born, Optik, pp. 209–218 (J. Springer, Berlin, 1933).
[12] H. Levine, Aperture diffraction near the geometrical optics limit, McGill Symp. Microw. Optics, pt. II, p. 211 (McGill Univ., Montreal, Canada, June 22 to 25, 1953).
[13] V. Fock, The distribution of currents induced by a plane wave on the surface of a conductor, J. Phys. (USSR) **10,** 130 (1946).
[14] I. P. Shkarofsky, M. P. Bachynski, and H. E. J. Neugebauer, Electromagnetic propagation over mountains, Can. IRE Conv. Record (Oct. 1958).
[15] J. R. Wait, Scattering of a plane wave from a circular dielectric cylinder at oblique incidence, Can. J. Phys. **33,** 189 (1955).
[16] J. R. Wait, Electromagnetic radiation from cylindrical structures (Pergamon Press, New York, N.Y., 1959).

(Paper 64D–64)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Characteristics of 488 Megacycles Per Second Radio Signals Reflected From the Moon[1]

B. C. Blevis and J. H. Chapman

(January 18, 1960)

Radio signals at 488 megacycles per second, received after reflection from the moon, have been studied using a continuous wave bistatic radar system located near Ottawa, Canada. These experiments were carried out during 1957 and consisted of monitoring the signals received on two orthogonal dipoles mounted at the focus of a 28 foot parabolic tracking antenna. The total mean received signal yielded an effective radar cross section of the moon at this frequency of 0.05 of the projected area. The libration fading as observed in the two orthogonal receiver channels was in synchronism, with a cross-correlation coefficient of 0.89. It was established that the signal is not depolarized significantly on reflection at the surface of the moon or in passing through the ionosphere. Rotation of the plane of polarization of the radio wave in the double passage through the ionosphere was observed to change by nearly 180 degrees over a six-hour period during quiet ionospheric conditions.

1. Introduction

Signals reflected from the moon have been used for studying propagation conditions along the path between the earth and the moon and for studying the radio reflection characteristics of the moon itself [1, 2, 4].[2] The study reported in this paper was begun primarily to investigate the rotation of the plane of polarization and the attenuation of UHF signals traversing the ionosphere. For this reason, plane polarized waves were transmitted from a tracking antenna, and the signals which were received after reflection from the moon were monitored by two independent receiving systems fed from identical orthogonal dipoles mounted in a second tracking antenna. By this means, the rotation of the plane of polarization could be observed and measured over a period of time.

The measurements were begun early in 1957 and were continued throughout the year, with various changes in equipment and procedure. Both long-pulse [2] and continuous-wave transmissions were used. For the latter measurements, it was necessary to separate the transmitter and receivers to prevent the very strong transmitted signal from saturating the receivers. For the cw measurements described in this paper, the separation was approximately 60 miles. Provided the direct signal was sufficiently low, the moon-reflected signal could be distinguished by means of its Doppler shift.

2. Experimental Method

The equipment used during these experiments is described briefly in the following paragraphs.

The transmitter was operated at a frequency of 488 Mc, with a cw power output of 10 kw. The transmitting and receiving antennas used 28-ft parabolic reflectors with equatorial mounts and drive systems. Small correction motors on each axis permitted uninterrupted tracking of the moon from moonrise to moonset. The transmitted radiation was essentially plane-polarized.

Two identical receiving systems with noise figures of about 7 db were fed from orthogonal dipole antennas mounted at the focus of the receiving parabolic reflector.

The transmitter frequency and the first two local oscillator injection frequencies of the receiver were controlled from similar crystal standards with an overall stability of the order of 1 part in 10^8. The receiver output at the second IF was then mixed with the signal from a stable VFO to give a third IF at a center frequency of 425 cps. The varying Doppler shift [2] due to the relative motion of the moon and the point of observation was eliminated by continual adjustment of the VFO to maintain a stable Lissajous pattern on an oscilloscope. In this way, the 425-cps output could be passed through a fixed narrow band (90 cps) filter. The signal was then recorded on magnetic tape for subsequent analysis, and also detected and displayed on paper chart recordings.

3. Signal Characteristics

3.1. Faraday Rotation

The Faraday rotation of the plane of polarization of a radio wave passing through the ionosphere has been described in detail in the literature [1]. A recording which exhibits the effects of Faraday rotation is given in figure 1. This record was made during a period of relatively quiet ionospheric conditions. The slow rotation of the plane of polarization of the signal during a 6-hr period is evident from a comparison of the two traces which come

[1] Contribution from Defence Research Telecommunications Establishment, Shirley Bay, Ottawa, Canada.

[2] Figures in brackets indicate the literature references at the end of this paper.

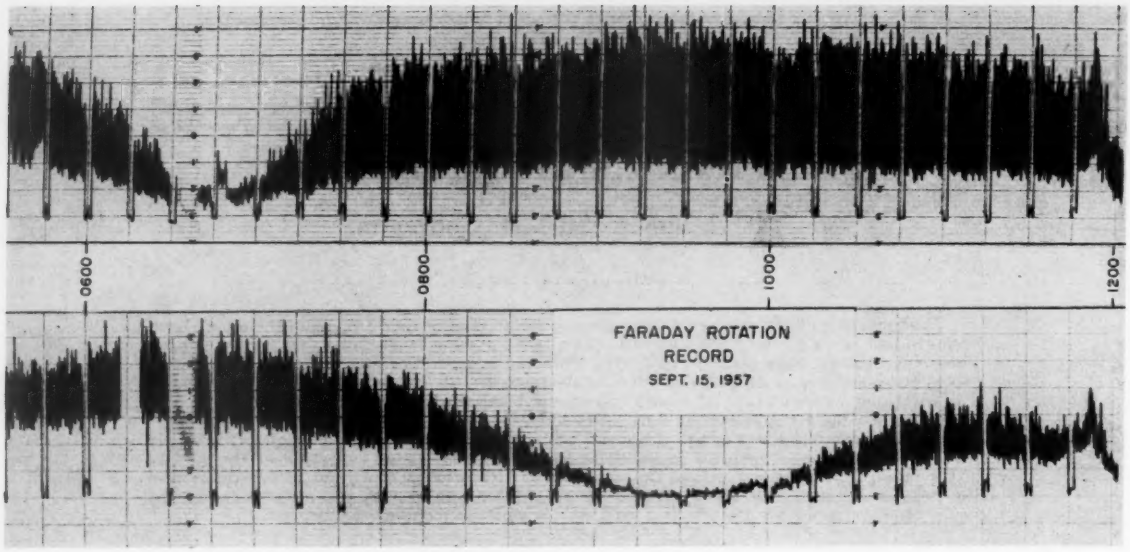


FIGURE 1. *Faraday rotation record.*

Comparison of the relative signal on two orthogonal polarizations over a 6-hr period on 15 September 1957.

from the two orthogonal receiver channels. The broadening of the record is due to the libration fading, as discussed below, which is partially smoothed by post-detection integration. Timing marks on the record were made by shutting the transmitter off every 15 min. Receiver noise can be seen during these periods.

Upper transit of the moon occurred just before the beginning of this record, and the moon was tracked down to the horizon. The increase of signal near the end is attributed to constructive interference between the direct and ground-reflected signals just before moonset.

This record was obtained during the morning when the total electron content of the ionosphere was increasing towards its maximum. In addition, the amount of rotation due to the increasing path length through the ionosphere was also increasing. It is reasonable to assume, therefore, that the rotation of the plane of polarization was monotonic, and hence a total change in rotation of almost 180 deg is indicated.

The difference in signal levels between the two orthogonal channels during periods when the received signal was nearly all in one of the two channels, or between maximums and minimums on a single channel, has been observed to be as great as 16 db. The isolation between the feed dipoles was 40 db when measured alone, and although this value was somewhat reduced when the dipoles were mounted in front of the reflector, the isolation was still greater than the measured differences between channels.

3.2. Libration Fading

Fading due to libration of the moon is generally deep and rapid, being as fast as several cycles per second during periods of maximum libration rate. This fading is due, of course, to the change in relative path lengths of signals scattered from different portions of the moon's surface.

A typical expanded recording covering a 35-sec period is shown in figure 2. The signal was about

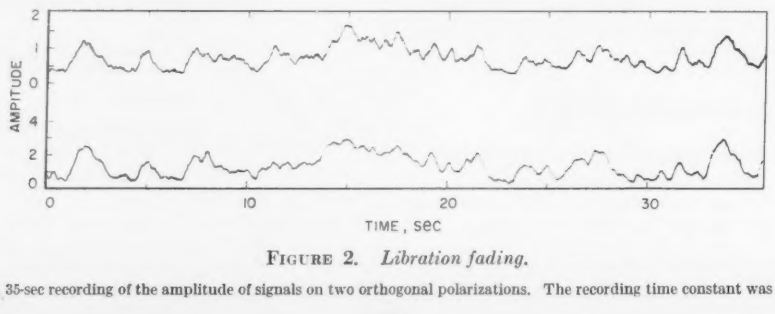


FIGURE 2. *Libration fading.*

A 35-sec recording of the amplitude of signals on two orthogonal polarizations. The recording time constant was 0.4 sec.

twice as strong in the upper channel but an adjustment of the recorder gain has compensated for this. The depth of fading will be limited by the effective bandwidth or amount of post-detection integration. However, in this example, with a time constant of 0.4 sec, the signal still fades into noise on occasion.

The signals are seen to be well correlated in the two receiving channels. Several analyses of this correlation have been performed. A typical example is illustrated by figure 3 which shows a scatter plot of the signal amplitude on the two channels at 0.8-sec intervals during a 1-min period. Machine calculations indicate a cross-correlation coefficient of 0.89.

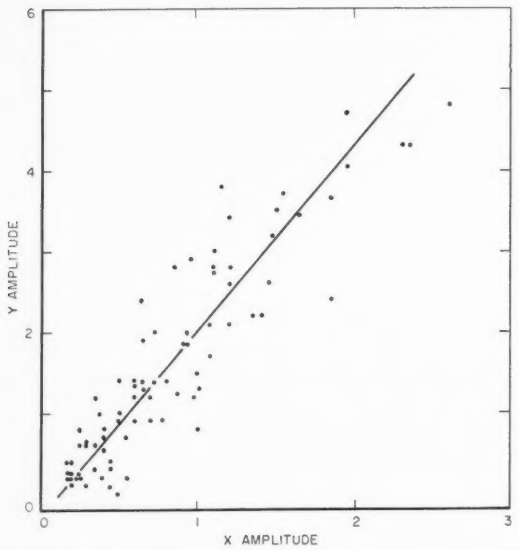


FIGURE 3. *Scatter plot of the relative signal amplitude on orthogonal polarizations.*

An amplitude of zero corresponds to the noise level.

The amplitude distribution of the signal on either channel approximates that of a Rayleigh distribution as can be seen by figure 4, which is a logarithmic plot of the probability density in amplitude. The amplitude is given in arbitrary units. These data were obtained by sampling the amplitude of a 5-min section of record at a sampling frequency of 1 kc.

From the analyses in this and the above section, one can conclude that there is little depolarization of the signal on reflection at the moon's surface or in its passage through the ionosphere. The plane polarized signal, which was transmitted to the moon, is received on the earth with the plane polarization largely preserved.

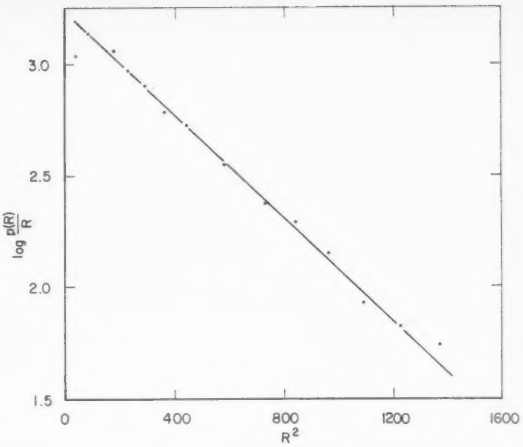


FIGURE 4. *Log plot of the probability density function.*

R is the signal amplitude on a single channel in arbitrary units, $p(R)$ its probability density.

3.3. Total Signal Power

A number of calculations have been made of the total signal power received by adding vectorially the signals in each channel. Assuming antenna gains of 30 db, the corresponding effective cross section of the moon is calculated to be of the order of 5×10^{11}m^2 or $0.05\pi a^2$, where a is the lunar radius. The uncertainty in this result is believed to be less than 3 db.

When these experiments began in February 1957, there were times when the total signal was unaccountably low for periods of the order of an hour or so. The apparent loss of signal was about 6 db or more. The times when this phenomenon occurred did not appear to correlate with the position of the moon above the horizon [3]. Investigations at the time were thought to rule out the possibility of any equipment malfunction.

Subsequently fairly exhaustive studies have been made in an attempt to repeat these observations under carefully controlled conditions using superior equipment [4]. The results have been negative, in that the total power was observed to remain sensibly constant within a decibel or so. It is concluded that either the physical conditions responsible for the earlier loss of signal were not duplicated during the subsequent tests, or that some still unknown equipment factor was responsible.

4. Conclusions

The Faraday rotation of the radio signal traversing the ionosphere in two directions has been observed to change by nearly 180 deg over a 6-hr period during quiet ionospheric conditions.

From an analysis of the rapid fading of the signal due to libration of the moon and of the difference in signal power between cross-polarized components,

it is concluded that the transmitted polarization is largely preserved on reflection from the moon's surface and that the surface is probably not deeply serrated and rough, at least at this wavelength (62 cm).

The effective value of the lunar radar cross section at 488 Mc, obtained by adding vectorially the mean signal on orthogonal polarizations, is of the order of $0.05\pi a^2$ as measured with a cw radar.

The results of more recent experiments carried out jointly by Lincoln Laboratory of M.I.T., the Stanford Research Institute, and the Defence Research Telecommunications Establishment have been published elsewhere [4].

The authors acknowledge the assistance of their colleagues, particularly F. D. Green, H. V. Serson, and E. A. Cameron, in making these measurements.

5. References

[1] I. C. Browne, J. V. Evans, J. K. Hargreaves, and W. A. S. Murray, Radio echoes from the moon, Proc. Phys. Soc. **B 69,** 901 (1956).

[2] B. C. Blevis, Ionospheric studies by the lunar radar technique, Nature **180,** 138 (1957).

[3] S. Rush and L. Colin, The effects on radio astronomical observations due to longitudinal propagation in the presence of field-aligned ionization, Proc. I.R.E. **46,** 356 (1958).

[4] J. C. James, L. E. Bird, R. P. Ingalls, M. L. Stone, J. W. B. Day, G. E. K. Lockwood, and R. I. Presnell, Observed characteristics of an ultra high frequency signal traversing an auroral disturbance, Nature **185,** 510 (1960).

(Paper 64D4–65)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

The Use of Polarization Fading of Satellite Signals To Study the Electron Content and Irregularities in the Ionosphere[1, 2]

C. Gordon Little and Robert S. Lawrence

(January 12, 1960)

A procedure is described for using the Faraday-rotation fading of a satellite radio signal to measure the ionospheric electron content per unit column up to the height of the satellite. At frequencies as low as 20 Mc the rotation of the plane of polarization cannot be assumed to be proportional to $\int NB \cos\theta\, dl$ along the line of sight. The simplifying assumptions implied by this expression are avoided, and full account is taken of ionospheric refraction, using the collision-free form of the Appleton-Hartree equation. Results based on observations of 1958 Delta 2 are presented. The subsatellite electron contents have been derived throughout the satellite passes for heights both above and below the F-peak; the latter compare well with values derived from simultaneous ionograms. The method also permits the study of large-scale irregularities in electron content. Such irregularities, having lateral dimensions of a few hundred kilometers and fractional deviations in subsatellite electron content of about 0.01, have been detected. Our observations suggest that satellite polarization studies offer important advantages over other methods of investigating these irregularities.

1. Introduction

The study of the ionosphere by means of the Faraday rotation of the plane of polarization of a radio wave has developed from the early moon-radar work of Murray and Hargreaves [1]. These and other authors have shown that, after making certain assumptions, one may write

$$\Phi = \frac{A}{f^2}\int_l NB\cos\theta\, dl \qquad (1)$$

where

Φ=the rotation of the plane of polarization,
A=a constant,
f=frequency of radio wave,
N=number density of electrons,
B=magnetic field strength,
θ=angle between magnetic field and direction of propagation,
dl=element of path along the line of sight l.

It has been customary to rewrite eq (1) in the form

$$\Phi = \frac{A}{f^2}\,[B\cos\theta\sec i]_{\text{avg}}\int_0^s N\, dh \qquad (2)$$

where i is the zenith angle of the line of sight, and a mean or "effective" value of $B\cos\theta\sec i$ is taken along the ionospheric part of the path.

Equation (2) shows that, within the limits of accuracy imposed by the approximations used in its derivation, an experimental measurement of Φ permits calculation of the electron content in a vertical column of the ionosphere. In practice, Φ is usually indeterminate by $n\pi$, and additional information is normally required before the value of $\int_0^s N dh$ can be derived unambiguously.

The approximations implicit in eq (2) are as follows:

(1) The observing frequency, f, is very much larger than the maximum plasma frequency, f_c, along the line of sight.

(2) The observing frequency is very much greater than the gyromagnetic frequency, f_H.

(3) The quasi-longitudinal approximation of the Appleton-Hartree equation may be used.

(4) The zenith distance, i, is small and the ionospheric thickness is very small compared with the radius of the earth. (These conditions restrict the variation of B, $\cos\theta$ and $\sec i$ along the line of sight and permit the use of an average value of each quantity along the ionospheric part of the path. They also minimize the errors which arise due to the neglect of ionospheric refraction. In the presence of such refraction, the radio energy does not traverse the single straight line assumed in eq (2); instead the ordinary and extraordinary waves traverse two different refracted paths.)

(5) The ionospheric electron density contours are spherically stratified.

In view of the above limitations, the moon-radar studies of ionospheric electron content have usually been made at frequencies in excess of 100 Mc and at fairly high angles of elevation.

[1] Presented at COSPAR First International Space Science Symposium, Nice, France, January 11 to 15, 1960.

[2] Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

The advent of satellite-borne radio transmitters has offered a new opportunity for Faraday-rotation studies of the ionosphere [2]. It is important to note that satellite studies, in principle at least, have the advantage that the electron content of the ionosphere can be measured over a wide range of latitudes from a single observing site, within the course of the few minutes duration of a satellite passage. In order to make full use of this advantage, however, it is necessary to remove the restriction to low zenith angles; it is also desirable to use frequencies relatively close to the ionospheric critical frequency in order to improve the sensitivity of the observations.

The failure of eq (2) at low frequencies and at large zenith angles is strikingly demonstrated by simultaneous observations of the 20- and 40-Mc signals from the Soviet satellites. Equation (2) predicts that the ratio of the Faraday rotation rates on the two frequencies shall be 4.0; in fact ratios of up to 5.0 have been observed, and on occasion, for a short period of time the sense of rotation has even been opposite on the two frequencies.

The purpose of this paper is to present a method of analysis whereby the assumptions implicit in eq (2) are eliminated. The method permits the deduction of accurate values of $\int_0^s N\,dh$ throughout a satellite pass, at frequencies as low as 20 Mc.

The main part of the paper opens with a description of the techniques used in recording the Faraday-rotation data. This is followed by an outline of the method of analysis; results of the application of the method to a low and a high pass of 1958 Delta 2 (Sputnik III) are given. The accuracy and limitations of the method are then examined. The paper concludes with a discussion of the large-scale ionospheric irregularities revealed by the analysis, and of the sensitivity and value of the Faraday-rotation technique for such studies.

2. Observational Methods

The satellite-observing facility at the Table Mesa field site (lat. 40°7.7′ N, long. 105°14.3′ W) of the Central Radio Propagation Laboratory was constructed in the spring of 1958. Interferometer, Doppler, and polarization measurements have been made each day since May 15, 1958, on the close passes of 1958 Delta 2, using a variety of antennas, baseline lengths, and baseline orientations.

The Faraday rotation records which form the observational data for this paper were taken using horizontal, linearly polarized 20-Mc dipoles mounted $\lambda/3$ above the ground. The antennas were connected via coaxial transmission lines to standard commercial communication receivers which were operated with the receiver AVC on. The AVC voltages were recorded on multichannel Sanborn pen recorders at chart speeds of 1 cm/sec. In the absence of AVC the receiver output would have varied quasi-sinusoidally as the plane of polarization rotated relative to the antenna. AVC was used in order to obtain an essentially logarithmic input-output receiver characteristic; the normal quasi-sinusoidal output waveform was thereby replaced by a waveform having broad maximums and very narrow, well-defined nulls. These nulls were often sufficiently well defined to permit their timing to an accuracy of 0.1 sec.

An alternative receiver output presentation has also been used. Two receivers having equal gain were connected to two similar antennas, one oriented N–S, the other E–W. The AVC outputs of the two receivers were subtracted and the resultant voltage recorded on the pen recorder. Since the receiver outputs for the two orthogonal antennas may be (approximately) expressed as log sin ωt and log cos ωt respectively (where ω=angular rate of rotation of the plane of polarization), their difference is equal to log (sin ωt/cos ωt)=log (tan ωt). The resultant waveform is nearly triangular, with equal positive and negative slopes and cusps at the null-times on each antenna. This form of presentation was found advantageous when the satellite signal was fading irregularly, or when the fading rate was so fast as to merge with the beeping of the satellite signal.

A third method, used occasionally, was to record the relative phase of the radio signals received on circularly polarized antennas of opposite sense of rotation. This method has the advantage of providing continuous information on the orientation of the plane of polarization, rather than information on null-times only. The technique was not found useful on 1958 Delta 2, primarily owing to the intermittent, beeping character of the signal.

3. Method of Analysis

Our method for the quantitative interpretation of the observed Faraday-rotation data is described below. The essential feature is the computation of the expected time variation of the specific rotatory power, $C(t)$, of the electrons along the line of sight, as the satellite moves along a known path through a model ionosphere. $C(t)$ is defined by

$$C(t)=\frac{\Phi_M(t)}{I_M(t)} \tag{3}$$

and is a measure of the efficiency of the electrons below the satellite in producing Faraday rotation. $\Phi_M(t)$ is the variation, with time, of the total Faraday rotation of the satellite signal as it traverses the model ionosphere en route to the observer, and $I_M(t)$ is the variation with time of the subsatellite electron content, $\int_0^s N\,dh$, in the model. $\Phi_M(t)$ is computed by a ray-tracing program, and $I_M(t)$ is readily obtained by an integration of the model ionosphere, hence $C(t)$ can be computed.

In order to derive $C(t)$ correctly, a moderately accurate ionospheric model is desirable, although as shown in section 5, high accuracy is not required. This fortunate feature arises from the fact that reasonable changes in the model affect $\Phi_M(t)$ and

$I_M(t)$ in almost the same proportion, and their ratio is therefore essentially unchanged.

Having determined $C(t)$ for the model in this manner, we make use of the observed values of Faraday rotation, $\Phi(t)$ to determine the true ionospheric electron content by writing

$$I(t)=\frac{\Phi(t)}{C(t)} \tag{4}$$

The observed values of $\Phi(t)$ are, in general, uncertain by an additive constant, $n\pi$; the manner by which this uncertainty may be reduced is discussed in section 4.

The method is now examined in greater detail; the reader will find it helpful to refer to figure 1 as each new step is discussed.

Orbital elements of the satellite were used to compute its position, relative to the observing station, at each of several selected times during the passage. The elements used were based on those of the Smithsonian Astrophysical Observatory or upon predictions made by the Vanguard Computing Center. The latter, being predictions, were much less accurate than the former but, if the time of closest approach is adjusted to agree with local observations, the remaining errors have been found to have slight effect on the derived values of I.

For heights up to F-peak, spherically stratified ionospheric models were used, based upon true-height analysis of local ionosonde data. The profiles above F-peak were obtained by assuming them to be those of a Chapman layer. This approach is a convenient one, as the Chapman function has the proper slope at the F-peak and asymptotically approaches a decaying exponential function at great heights. After fitting the function at F-peak, only one adjustable parameter remains, the asymptotic exponential decay constant. We shall refer to this constant as the "electron scale height." It should be noted that the Chapman function is here used only as a convenient mathematical model, without any regard to its physical basis. If the values of I obtained from a satellite which is above the F-peak differ significantly from the corresponding values of I_M, the electron scale height may be adjusted until agreement is reached.

The assumption of spherical stratification can be removed by the inclusion of a horizontal gradient in the ionosphere model. One method for the determination of such a gradient from spaced ionograms is illustrated in figure 2. Here values of f_0F2 have been plotted from simultaneous observations over a wide area, and contours of f_0F2 drawn. The values of critical frequency thus deduced along the satellite track (and assumed to be at the same true height) provide factors by which the electron-density profile at Boulder must be multiplied to approximate the actual ionosphere along the track. The necessity for including horizontal gradients adds no particular difficulty to the ray-tracing procedure, but it provides another adjustable parameter which must be determined, and thus it weakens the determination of electron scale height above the F-peak.

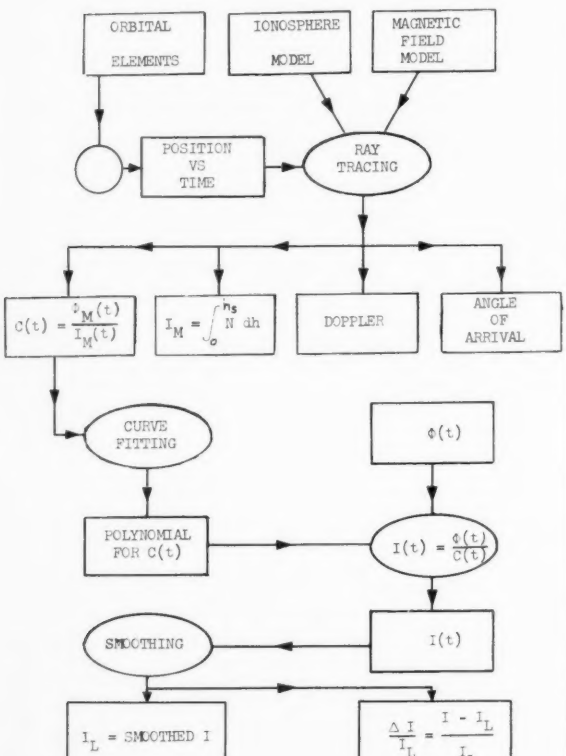


FIGURE 1. *Method of analysis of satellite Faraday rotation records.*

The circles or ovals represent processes carried out with a digital computer; the rectangular boxes represent input or output data for these processes.

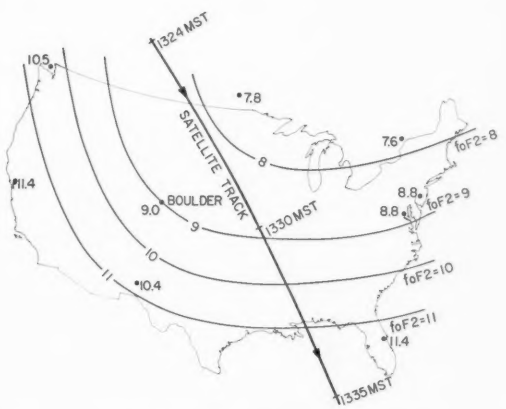


FIGURE 2. *Geographical variation of f_0F2, used to estimate the horizontal ionospheric gradient, September 17, 1958, 1330 m.s.t.*

The geomagnetic field is commonly represented by the field of a dipole of appropriate strength and orientation at the center of the earth. When, as in the present application, the field is needed over only a limited portion of the earth, a significant improvement can be made by adjusting the strength and orientation of the dipole to best fit the observed field in that region. Such a "modified dipole" field has been used for some of the results presented here. A more accurate, but much more laborious method is to add higher-order terms to the spherical-harmonic representation of the field. This method has been used with terms up to order six, but only for comparison with the modified dipole results because it involves an impractically long calculation. When used in ray tracing, the 48-term polynomial for the magnetic potential must be solved four times to calculate the vector magnetic field. For each position of the satellite, this entire operation is repeated at dozens of points along each ray, and must be done not only for the final rays but for all the trial rays which fail to connect the satellite and the observer. The relative accuracy of results based on the modified dipole field and the complete spherical-harmonic representation is discussed in a later section.

Once the satellite motion is known and models have been chosen for the ionospheric electron-density profile and for the geomagnetic field, the specific rotatory power, $C(t)$, of the ionosphere can be obtained by ray tracing. Since C depends upon the direction of the geomagnetic field relative to the ray and on the zenith angle of the ray, it obviously varies with time during the passage of the satellite. In obtaining Φ_M, and hence $C(t)$, we have used a variety of ray-tracing programs.

The simplest program we have used evaluates $\int_o^s N B \cos\theta \sec i \, dh$ along the straight line from the observer to the satellite. This is an improvement over the use of eq (2) in that it eliminates the assumption of a thin ionosphere and, to some extent, permits the use of larger zenith angles and horizontal ionospheric gradients. The high-frequency and quasi-longitudinal approximations remain, however, and the effects of refraction are ignored.

A second program, designed to overcome the difficulties of the high-frequency approximation, evaluates the integral $2\pi/\lambda \int_0^s (\mu_0 - \mu_e)\, dl$ along the straight line from observer to satellite. Here μ_o and μ_e are the local refractive indices of the ionosphere for ordinary and extraordinary waves traveling along the line of sight. Half of this integrated difference between the ordinary and extraordinary phase path lengths gives Φ_M. I_M is determined by simple integration of $N\, dh$ along the straight line. Since this second program uses the complete collision-free form of the Appleton-Hartree equation (see for example, ref [3]) to evaluate μ_0 and μ_e, the high-frequency approximation is removed. For nearly transverse propagation the Faraday rotation is critically dependent upon θ, so it is meaningless to speak of removing the quasi-longitudinal approximation as long as refraction is neglected.

A third program, involving considerably more computer time, includes most of the effects of refraction. This program evaluates $2\pi/\lambda \int_0^s (\mu_0 - \mu_e)\, dl$ along the refracted path which the radio ray would follow in the absence of the geomagnetic field. This program assumes that the ordinary and extraordinary rays traverse the same refracted path. It works well in the quasi-longitudinal part of the sky but is quite unsatisfactory near the quasi-transverse region, where the Faraday rotation is critically θ-dependent.

A final program has been developed for use in this critical portion of the sky. It is a complete ray-tracing program which evaluates separately the phase paths, P_o and P_e, of the ordinary and extraordinary rays, along their independent refracted paths. This full ray-tracing is accomplished by using Snell's law at the interfaces of a series of elementary spherical shells, each considered to possess constant refractive index. Referring to figure 3, let O_i be the intersection of the ray with the interface between spherical shells i and $i+1$. $\mathbf{N_i}$ is the unit vector normal to the interface and $\mathbf{W_i}$ is the unit vector in the direction of the wave normal in shell i. $\mathbf{B_i}$ is the geomagnetic field. The electron densities, N_i and N_{i+1}, are known in each shell. The refractive index, μ_i, is known, but μ_{i+1} is not known since it depends upon θ_{i+1}. A first approximation to μ_{i+1} may be obtained by setting $\theta_{i+1}=\theta_i$. Snell's law then gives a second approximation to θ_{i+1}, and hence to μ_{i+1}. Two or three iterations are required to determine μ_{i+1} and θ_{i+1} with sufficient accuracy, and thus to determine the direction $\mathbf{W_{i+1}}$ of the new wave normal. As discussed by Bremmer [4], the new ray, $\mathbf{R_{i+1}}$, does not coincide with $\mathbf{W_{i+1}}$ but is separated from it by an angle α. $\mathbf{R_{i+1}}$ lies in the plane of $\mathbf{N}$ and $\mathbf{B}$ and so is not necessarily in the plane of incidence. An extension of $\mathbf{R_{i+1}}$ through the shell defines O_{i+1} at the second interface and completes the ray tracing through shell $i+1$.

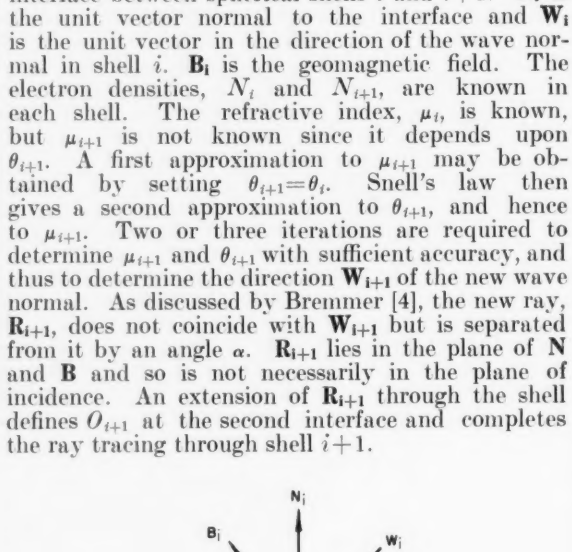


FIGURE 3. *Detail of ray tracing through one spherical shell.*

When a ray has been traced from the satellite to the ground it will not, in general, arrive at the observer. It is then necessary to make a correction to the direction of departure from the satellite, and trace another ray. Four to eight such attempts, depending upon the zenith angle, are usually sufficient to discover the ray which, with sufficient accuracy, connects the satellite to the observer. The accuracy required is such that the small difference in phase paths, P_o–P_e, is known to an accuracy of one percent. (This typically requires the determination of each phase path length to about one part in 10^6.) Shell thicknesses as small as 10 km are generally required. In extreme cases, the IBM 650 computer may require from 30 to 60 min to complete the process for a single position of the satellite. Accordingly, this full ray-tracing program is used only near the quasi-transverse regions of the sky, or to evaluate the accuracy of the faster, more approximate programs.

The ray-tracing procedure just described is strictly accurate only when the surfaces of constant refractive index coincide with the spherical shells. In fact, it has proved to be sufficiently accurate in the case of all reasonable horizontal gradients in the ionosphere. In these cases, the electron density of each shell must be given in terms of the geographical location of the ray, but comparison with a more rigorous method has shown that the interfaces between shells may still be considered to be normal to the radius vector from the center of the earth. A detailed description of the ray-tracing program is in preparation.

Byproducts of the ray-tracing procedure are the Doppler frequency shift and the angle of arrival. The observed Doppler shift depends not only upon the radial velocity of the satellite, but also upon the local electron density and upon the rate of change of integrated electron density along the ray path. These ionospheric effects are automatically included in the results of the ray tracing. Similarly, the ray-tracing program evaluates the difference in angle of arrival between the ordinary and extraordinary rays, the total bending of each, and the spatial separation between each ray path and the line of sight.

It is clearly impractical to evaluate $C(t)$ at the time of each observed Faraday-rotation null. Since its variation depends only on the geometrical position of the satellite and upon the geomagnetic field along the ray, $C(t)$ is a smoothly varying function of time and can be represented very well by a simple polynomial. Typically, $C(t)$ has been evaluated by ray tracing at nine equally spaced times throughout the satellite passage, and the method described by Anderson [5] has been used to fit, by least squares, a fifth-degree polynomial to these nine points. In every case the polynomial has agreed with the nine given values to within a small fraction of one percent.

Once a polynomial expression has been obtained for $C(t)$, any given value of $\Phi(t)$ is readily converted to a corresponding value of I by means of eq (4). Successive Faraday-rotation nulls which appear on the recordings of the satellite signal correspond to values of Φ which differ by π radians. It is necessary only to observe the times at which such nulls occur, and then to assign to some one null an assumed value of Φ, in order to obtain a list of values of $\Phi(t)$. The initial value of Φ is usually selected to be equal to the corresponding Φ_M from the ray-tracing program. A method of improving this initial estimate is demonstrated in the next section.

In some records the direction of rotation of the plane of polarization is found to reverse, i.e., $d\Phi/dt$ changes sign. It is important to recognize such events on the record, and this can be done most easily by noting the abrupt change in relative phase of the fading on two orthogonal linearly-polarized antennas.

Details of the relatively small fluctuations which occur in $I(t)$ become evident only by separating them from the gross value of I. $I(t)$ was therefore smoothed by a running average of suitable duration, to give $I_L(t)$. The duration of this smoothing interval was chosen to equal the time required for the portion of the ray which is 300 km above the ground to move through a distance of 400 km. The exact duration of this interval is therefore a function of the height of the satellite and of its velocity. The fractional irregularities in $I(t)$ are given by

$$\frac{\Delta I}{I_L}=\frac{I-I_L}{I_L}$$

This process of taking the difference from a smoothed value is in fact a spatial filter. It has the property of rejecting all Fourier components with wavelengths much above 400 km and has relatively little effect upon the amplitude of components smaller than the smoothing interval.

4. Some Preliminary Results of the Method

Figure 4 is a plot of the derived subsatellite electron content of the ionosphere during the northbound pass number 1073 of 1958 Delta 2. This

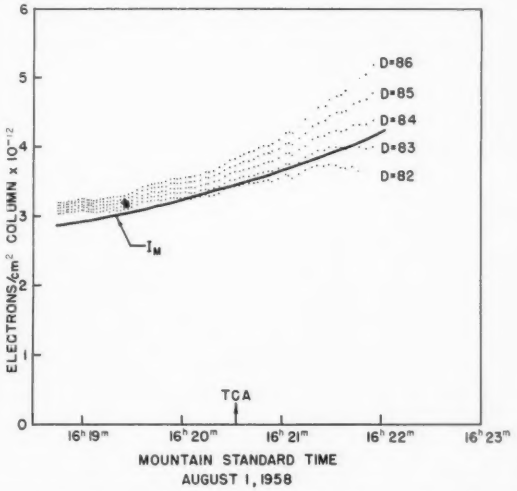


FIGURE 4. *Comparison of* I_M *with Faraday rotation data for five values of* D.

pass took place on the afternoon of August 1, 1958; perigee occurred with the satellite near the southern horizon, so the satellite was rising slowly as it passed almost overhead at a height of about 232 km.

The model of the ionosphere used in the analysis of figure 4 was obtained by a linear interpolation between the true-height electron profiles deduced from ionograms taken at Boulder at 1600 and 1700 m.s.t. The available orbital elements have been adjusted to give the correct time of closest approach to better than 1 sec. In the ray tracing for this particular pass, the 48-term expansion of the magnetic field was used, in order to obtain maximum accuracy. The derivation of $C(t)$ was made as described in section 3 above, using the full ray tracing program. The ray tracing program permits an estimate of the total number, D, of half rotations existing at the time of the first Faraday rotation null; in order to further refine this value, a family of curves differing only in value of D was plotted. The derivation of the most probable value of D was made as follows.

In a fixed, spherically stratified ionosphere, the rate of change of the subsatellite electron content is

$$\frac{dI}{dt}=N_s\frac{dh}{dt},$$

where N_s is the electron number density at the satellite, and dh/dt is the rate of increase of height of the satellite. For given orbital elements, dh/dt is known and it is therefore possible to compute observational values of N_s for each value of D as a function of time throughout the satellite passage. When this is done, very large differences in N_s are obtained for different values of D, particularly near the quasi-transverse region of the sky. Thus, the values of N_s deduced for the last nine points of each D curve of figure 4 vary from negative values ($D=82$), to approximately zero ($D=83$), a value approximately equal to the expected N_s ($D=84$) and values roughly 1½ and 2 times greater than expected ($D=85$ and 86 respectively). For this reason, one can state the value of D almost certainly lies in the range 83 to 85.

Taking a value of $D=84$, one obtains a value of $I=3.62\times10^{12}$ electrons per square centimeter column at the time of closest approach. The corresponding value of I deduced from the model (itself based on the ionosonde data) is 3.46×10^{12}, a discrepancy of 4.5 percent. The difference between I and I_M is roughly constant throughout the satellite pass, and can very readily be explained by the effect of errors in D, in the magnetic field, in the ionospheric model used, in the assumed height of the satellite, and/or in the neglect of satellite spin and ionospheric gradients.

One striking feature of the $I(t)$ values derived in this way is that the values do not lie on a smooth curve. Figures 5, 6, and 7 are plots of $\Delta I/I_L$ for three passes at heights of the order 230, 300, and 1250 km. As indicated in section 3, the irregularities are relative to running smoothed values of I, the smoothing interval being adjusted for each pass to correspond to a 400-km running average at a height of 300 km.

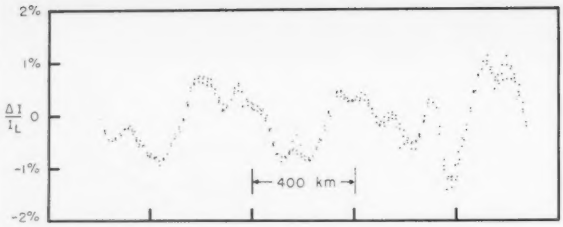


FIGURE 5. *Irregularities in subsatellite ionospheric electron content, 1620 m.s.t., August 1, 1958, as observed on three independent equipments.*

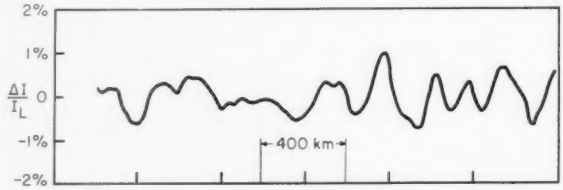


FIGURE 6. *Irregularities in subsatellite ionospheric electron content, 0830 m.s.t., September 4, 1958.*

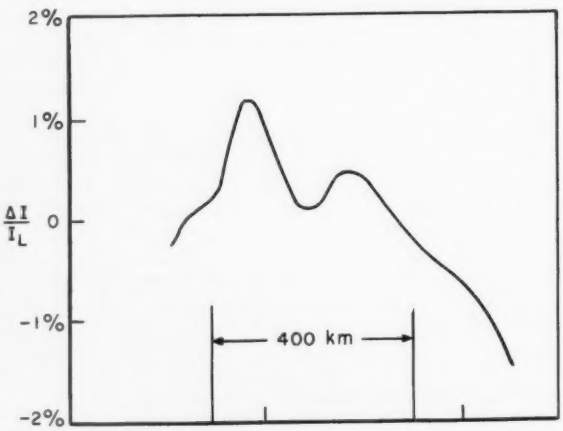


FIGURE 7. *Irregularities in subsatellite ionospheric electron content, 1330 m.s.t., September 17, 1958.*

Figure 5 is a mass plot of the values of $\Delta I/I_L$ resulting from the Faraday-rotation data taken simultaneously on three separate receivers coupled respectively to dipoles oriented N–S, E–W, and from NW to SE. The excellent agreement of the points strongly indicates that the irregularities in polarization rotation which gave rise to the irregularities in I are real and are not due to such effects as scaling errors, or satellite signals scattered by objects near the antennas.

Figures 6 and 7 are of interest in that the irregularities observed on two very different passes are similar in fractional deviation and in size. Figure 6 refers to a south-north pass at a height of 290 km on September 4, 1958, and figure 7 to a 1,250-km height, north-south pass on September 17, 1958. Also of interest is that a pass observed by Dr. Owen Garriott of Stanford University, and analyzed by our method, shows similar irregularities. This fact would appear to confirm that the phenomenon is real, and is not introduced by our equipment.

5. Discussion of Accuracy

The accuracy with which I, the total electron content per unit column to the height of the satellite, can be derived by the Faraday-rotation method depends upon many factors, although the accuracy of measurement of the irregularities in I is primarily limited by scaling errors. The magnitude of typical scaling errors is illustrated by the scatter of the points in figure 5. It is quite small compared to the magnitude of the ionospheric irregularities themselves.

In contrast to angle-of-arrival studies, the Faraday method does not require extremely accurate knowledge of the position of the satellite. For example, the low pass illustrated in figure 4 was analyzed independently with two different sets of orbital elements. These were Vanguard predictions issued respectively 1 day before and 6 days after the passage of the satellite. They gave apparent minimum zenith angles at Boulder of 1° and 10°, and heights at closest approach of 236 and 232 km, respectively. In each case an arbitrary time correction of nearly a minute was needed to obtain agreement with the observed time of closest approach at Boulder. Despite the definite disagreement between the two sets of elements, the resulting values of I differed by less than 3 percent in all parts of the sky except the far north, where transverse propagation and the resulting critical dependence upon θ increased the difference to 15 percent.

It has already been mentioned that $C(t)$, and therefore $I(t)$, are nearly independent of the ionosphere model used in the analysis. The extent to which this is true is illustrated in figure 8, where the analysis has been carried through for two ionospheric models differing by a factor of two in electron density at each height. The resulting values of I differ by only 3 percent near the zenith, and by less than 10 percent at the ends of the recorded passage. As pointed out earlier, this fortunate behavior results from the fact that a change in I_M causes a nearly proportional change in Φ_M.

In figure 9 we demonstrate, using the passage of August 1, 1958, the importance of utilizing an accurate model for the geomagnetic field. The lowest curve shows the $I(t)$ which results from the use of the complete spherical-harmonic expansion of the field. The Gaussian coefficients used here are those published by Finch and Leaton [6]. In the absence of better knowledge of local geomagnetic anomalies,

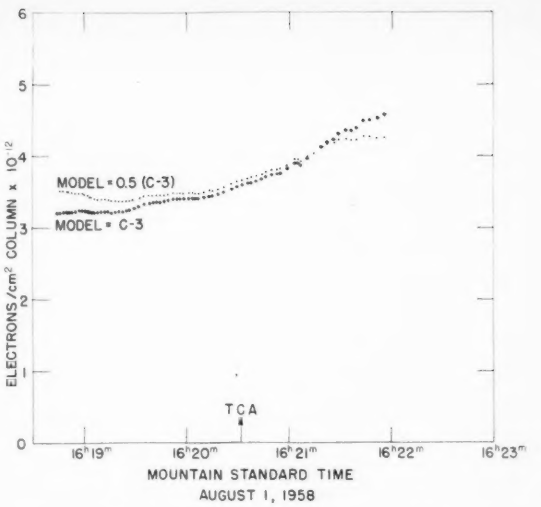


FIGURE 8. *Effect of different ionospheric models upon the derived subsatellite electron contents.*
TCA = time of closest approach.

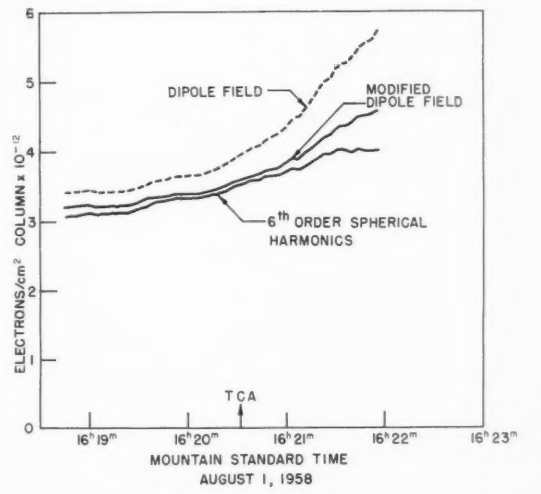


FIGURE 9. *Effect of different magnetic field models upon the derived subsatellite electron contents.*

including ionospheric currents, these coefficients represent the most accurate means we have available for estimating the geomagnetic field in the ionosphere. Figure 9 shows, for comparison, the result of using the standard centered dipole field, and the field of our "modified dipole." The modified dipole, adjusted to agree with the field at Boulder, removes about two-thirds of the total error caused by the standard centered dipole. Although an accurate magnetic model is necessary for estimating I, there is no such requirement for $\Delta I/I_L$. The modified

dipole or, if we except the transverse-propagation region near the northern horizon, even the standard dipole, would suffice for the study of the irregularities, since these are measured relative to the corresponding smoothed values.

Figure 10 illustrates the importance of including ionospheric refraction. Here the correct value of $I(t)$, including all refractive effects, is compared with another which neglects only refraction. This latter curve results from evaluation of $(2\pi/\lambda)\int_0^s(\mu_0-\mu_e)dl$ along the straight line joining the observer and the satellite. The complete Appleton-Hartree equation (excluding absorption) is used to evaluate μ_0 and μ_e, so the high-frequency approximation is avoided. As might be expected, neglect of refraction has no effect near the zenith but becomes serious near the horizon. At the northern horizon the effect is more pronounced than in the south, and it is of opposite sign. These differences result from the dependence of Faraday rotation upon cos θ, a factor which varies rapidly near the northern horizon.

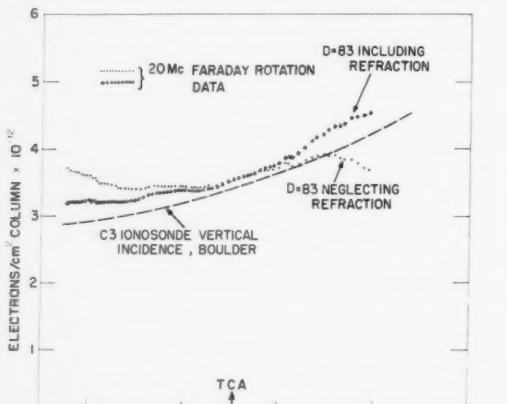


FIGURE 10. *Effect of including and omitting ionospheric refraction upon the derived subsatellite electron contents.*

The preceding discussion of accuracy assumes the existence of a spherically stratified ionosphere upon which irregular fluctuations may be superposed. Most of the remarks, particularly those concerning $\Delta I/I_L$ continue to be valid in the presence of systematic horizontal gradients, but there is one notable exception. The determination of D, and therefore of the absolute value of I, may be seriously influenced by such gradients. The implications of this effect are discussed in some detail in the next section.

6. Discussion and Interpretation of the Results

The discussion and interpretation of the results of analyses presented in earlier sections may be divided into (a) those dealing with the $I(t)$ values directly and (b) those dealing with the observed irregularities in $I(t)$.

6.1. Total Content

In view of the excellent agreement (for passes below F_{max}) between the derived $I(t)$ curves and those based on the corresponding ionosonde data, it may be claimed that the method is an accurate one and is capable of deriving the subsatellite electron content with relatively little uncertainty. It should be emphasized that the agreement shown in figure 4 is in no way exceptional and in fact is not as good as that obtained on at least one other pass. This excellent agreement may be taken as confirmation of the absence of any deep valley in the electron profiles between the E and the F regions, since the ionosonde profiles were derived assuming the absence of such valleys.

The main limitation to the accuracy is in the determination of D by the comparison of the slope of the model $I_M(t)$ and the derived $I(t)$ curves. This method is unfortunately sensitive to horizontal gradients in the electron content, and without prior knowledge of these gradients, it is apparent that significant errors could be made in selecting D. It should be realized that this difficulty applies also to the Faraday "fading-rate" method used by several authors [7, 8, 9].

Investigation of the change in dI/dt with D shows that a gradient of $\int_0^s N\,dh$ of 2 percent per 100 km along the pass would change the fading rate in the zenith, and hence the deduced value of $\int_0^s N\,dh$, by about 12 percent. Such gradients of $\int_0^s N\,dh$ must be expected, since even if they existed from pole to equator they would produce only a four-fold variation of total ionization. Over such distances the f_0F2 values typically vary by a factor of 3, implying, for low passes at least, a nine-fold variation in subsatellite electron content.

It should also be realized that, once D has been determined unambiguously, (for example, by simultaneous polarization measurements on two closely spaced frequencies) then the present method will become a sensitive one for investigating ionospheric gradients.

When the satellite is at a height well above the F-peak it is not possible to base the model ionosphere upon ionograms alone. Some assumption must be made concerning the shape of the upper portion of the electron density profile. As mentioned previously, the Chapman function is a convenient one to use. Its adjustable parameter, the electron scale height, must be given some arbitrary trial value. If there is reason to suspect that horizontal gradients may exist, it is, in addition, necessary to make some assumptions concerning them. Figure 11 illustrates the result of such an analysis of the record from a high satellite. The passage is the one shown in figure 2, from which it is obvious that some horizontal gradient should be included in the ionospheric model.

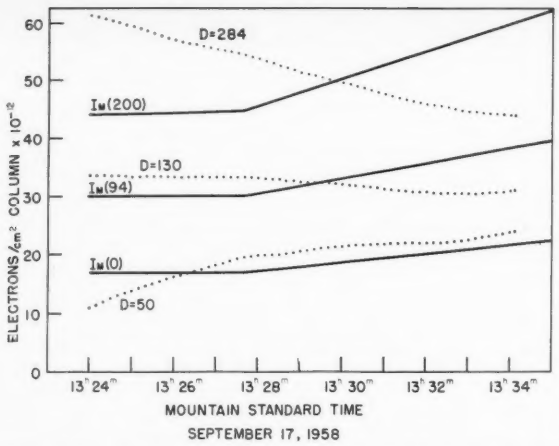


FIGURE 11. *Diagram illustrating the failure of the experimental data to fit the nonspherical ionosphere deduced from the* f_0F2 *data of figure 2.*

The simplest assumption is that the electron density varies with geographical position in the same way at all heights, and that f_0F2 can be used as an indicator of this variation. Acting upon this assumption we obtain for $I_M(t)$ the three curves shown in figure 11, for assumed values of electron scale height at Boulder of 200 km, 94 km, and zero. Zero scale height means, of course, that no electrons exist above the F-peak. The other three curves in the figure illustrate the resulting values of $I(t)$ for various values of D. The $I(t)$ curves shown were calculated with the ionospheric model which has an electron scale height of 200 km; the corresponding $I(t)$ curves for the other models are essentially the same.

A self-consistent picture demands that $I(t)$ agree with $I_M(t)$, and in this case it is obvious from figure 11 that no combination of scale height and D can produce such agreement. A change must be made in the ionospheric model, in such a sense as to reduce the rate of increase of subsatellite electron content with time. The f_0F2 values increased markedly along the track, and we have assumed an equivalent increase in electron density at all heights. Figure 11 shows that this assumption cannot be valid, and that the gradients must become relatively smaller above F-peak. In other words, the electron density at levels well above the F-peak must be less dependent upon geographical location than at F-peak.

The quantitative introduction of a new parameter, the rate of change of the horizontal gradient with height, is clearly unwarranted without additional data, though it may become possible to estimate this parameter after a large number of satellite records have been analyzed.

6.2. Irregularities

The irregularities in electron content shown in figures 5, 6, and 7 are of considerable interest, primarily because of their large size. These irregularities are, of course, derived from polarization-rotation data, and before discussing their significance one should first consider the validity of their derivation. As discussed in section 4, the excellent agreement between the values of $\Delta I/I$ derived from three completely independent observing systems precludes the possibility that the irregularities are introduced by scaling errors, or by scattered signals. One is therefore forced to conclude that the incident polarization vector did indeed rotate irregularly. Such irregular rotation could be introduced by an irregular motion of the satellite, by spatial or temporal irregularities in the magnetic field, and by spatial or temporal irregularities in the electron content of the ionosphere.

The effect of the rotation of the satellite upon the observed polarization is dependent upon the nature of the spin. A rapid, irregular tumble of the satellite can be ruled out, owing to the almost complete absence of external forces. The nonexistence of such an irregular motion was demonstrated by certain of the radio observations, which showed a regular sequence of nulls, superimposed upon the normal Faraday rotation, and occurring simultaneously on all polarizations. These nulls are interpreted as due to satellite rotation, the signal fading each time the rotating dipole was directed toward the observer. The time interval between these nulls was approximately 45 sec, implying a full rotation period of about 90 sec.

The effect of a regular rotation of the satellite upon the observed polarization is dependent upon the relationship between the orientations of the spin axis and the satellite antenna. In the case where the spin axis is parallel to the dipole, no effect will be observed. For orthogonal spin and dipole axes, the satellite spin will add to or subtract from the Faraday rotation, the correction reversing in sign as the observer passes through the plane of the rotating dipole. In the case where the angle between the spin and the dipole axes is neither zero nor 90°, it is convenient to differentiate between two conditions (a) where the observer lies inside the cone formed by the rotating antenna and (b) where the observer lies outside this cone. In general, there will be a transition between case (a) and case (b) as the satellite moves across the sky. In the region of sky defined by case (a), the satellite spin will add to or subtract from the Faraday rotation; for case (b) the observer will see the normal Faraday rotation plus a sinusoidal rocking of the polarization vector, of amplitude equal to the angle between the spin and the dipole axes, and period equal to the full rotation period of the satellite. In the former case the satellite rotation will introduce a cumulative error to the derived values of $I(t)$; for the latter case, quasi-sinusoidal oscillations of $I(t)$ of amplitude not greater than that introduced by a $\pm 90°$ shift in polarization rotation will be observed.

The derived irregularities of figures 5, 6, and 7 are not consistent with the suggestion that they are due to a regular satellite rotation. The irregularities are not sinusoidal, as would have been expected if they were due to satellite rotation. Instead of a period of about 90 sec, the irregularities

of figures 5 and 6 had periods of the order 40 to 50 sec, while those of figure 7 had periods considerably greater than 100 sec. Although the amplitude of the polarization flutter responsible for the irregularities of figures 5 and 6 was less than 90°, and was therefore consistent with the satellite-rotation hypothesis, the polarization irregularities observed during the pass of figure 7 were several times too large to be due to a regular satellite rotation. In addition, the amplitude of the irregularities in $\Delta I/I_L$, if due to satellite rotation, would be inversely proportional to the total amount of Faraday rotation existing at any time. A four- or five-fold increase in the amplitude of the irregularities should therefore have been observed in figures 5 and 6 as the satellite moved from the southern horizon toward the quasi-transverse region; no such increase was observed. For these reasons, therefore, we conclude that the observed irregularities in polarization angle were not due to satellite rotation.

Owing to the magnitude of the phenomenon (roughly ± 1 percent variation in total amount of polarization rotation), it does not seem possible that the observed irregularities could be due to temporal variations in the magnetic field or in the subsatellite electron content. For the same reason, it does not seem likely that they were due to spatial irregularities in the magnetic field. For this to be the case, the longitudinal component of the geomagnetic field would have had to vary irregularly, relative to our magnetic model, by roughly ± 1 percent in distances of the order of 300 km. Since maps of the surface magnetic field do not show such irregularities in the central United States, it would seem unlikely that they could exist at heights of the order of several hundred kilometers. It should be possible to test for such an effect by correlating the irregularities deduced from similar satellite passes, but as yet this test has not been carried out due to lack of fully analyzed data.

There remains one final alternative explanation of the deduced irregularities. The values of $C(t)$ are derived by a curve-fitting process, and any irregular variations in the computed $C(t)$ relative to the true $C(t)$ for a spherically stratified ionosphere would show up as irregularities in the electron content. The curve fitting process has therefore been tested. An analytical expression having a variation similar to that of $C(t)$ was obtained, and nine points along it were determined. The curve fitting program was then used to derive the best fitting 5th degree polynomial for these nine points. The agreement between values derived from the polynomial expression and from the original analytical expression was always within 0.1 percent, indicating that irregularities in the derived $C(t)$ curves were not responsible for the deduced variations in electron content.

On the basis of the above discussion, we conclude that the observed irregularities are correctly ascribed to variations in subsatellite electron content. We proceed now to discuss the properties of the irregularities; to relate our observations to other observations of large scale ionospheric irregularities, and to evaluate the Faraday-rotation method relative to other methods for observing such irregularities.

The following points may be made concerning these observed variations in subsatellite electron content:

(a) They have spatial extent of the order 300 km.

(b) Their intensity is such as to modulate the subsatellite electron content by roughly ± 1 percent.

(c) They are not limited to a narrow height range (otherwise the fractional deviation would be considerably less for high passes than for low passes).

(d) On occasion, at least, they were present for heights as low as 230 km.

(e) They are of common daytime occurrence at latitude 40 deg. The four passes analyzed in detail to date all show their presence; a cursory examination of the observational data suggests that they are present on many, perhaps even the majority, of the records.

(f) Their intensity apparently is limited to maximum values of fractional deviation of the order of a few percent.

(g) There is no evidence as yet concerning their diurnal, seasonal, or magnetic storm dependence, though information concerning these is presumably available in the numerous, as yet unanalyzed, records.

Several authors have given experimental evidence of the existence of large-scale ionospheric irregularities. The observations of large-scale traveling disturbances in the ionosphere by Munro [10] suggest that the irregularities he observed are related to our own observations. From Australian studies of the virtual height of reflection of HF radio waves, he was able to deduce the frequent occurrence, during daytime hours, of traveling disturbances in the F region having dimensions up to several hundred kilometers. Bramley and Ross [11] using angle of arrival measurements, and Whitehead [12], and Jones et al. [13] from studies of the "bursts" in amplitude of ground-based transmissions have been able to confirm the existence of similar irregularities in the northern hemisphere. In general, however, quantitative data on the intensity of the irregularities is lacking, and it is therefore not possible to be certain that the irregularities we observe are of the Munro type; in particular, our observations to date do not permit the study of their motion.

Evidence of large-scale ionospheric irregularities has also been obtained by Vitkevitch and Kokurin [14, 15] using radio astronomical sources. These authors made measurements of the angle of arrival of radio waves from the celestial radio sources and found evidence of slow oscillations in the apparent source position relative to the true source position. They estimate that the irregularities in total electron content of the ionosphere have lateral dimensions of the order 300 to 400 km and intensity of up to several percent, in good agreement with our observations of the variations in subsatellite electron content.

Unlike the satellite data, the nonsatellite methods described above have not given more than rough

estimates of the irregularities in ionospheric electron content; and in particular the nonsatellite methods do not readily permit the rapid study of the irregularities over wide geographical areas from a single site. It is of interest to consider whether the other parameters of a satellite radio wave (amplitude, frequency, and direction of arrival) may also be useful in the study of these large-scale irregularities. Examination shows that these alternative methods are unlikely to prove effective. The recordings of amplitude "bursts" of satellite signals, analogous to the amplitude studies of Whitehead, and Jones et al., would not appear to be a suitable method, owing to difficulties arising from the very great variations of signal strength due to the changing satellite range, and the changing efficiencies of the polarized transmitting and receiving antennas along the (moving) line of sight. A second effect of the ionospheric irregularities is that they will modify the ionospheric distortion of the free-space Doppler curve of a moving satellite. Any Doppler method for studying these large-scale irregularities, however, suffers from the fact that the Doppler technique, which is based on the rate of change of phase, is inherently insensitive to weak, large-scale irregularities. Thus, it can be shown that the irregularities of figures 5 and 6 would produce an irregular distortion of the Doppler curve of magnitude about 0.1 cps and period roughly 40 sec. Since the rate of change of the 20-Mc signal is typically several cycles per second per second, it would be difficult to measure such distortions accurately. The angle-of-arrival measurements, which are based on the rate of change of phase across the ground, are also insensitive to weak, large-scale irregularities. In this case, it can be shown for the passes of figures 5 and 6 that the apparent position of the radio satellite would have varied irregularly, relative to its true position, by about 1 min of arc and with period about 40 sec. Such irregular refractions would be difficult to measure accurately in an object which is moving through as much as 100 min of arc per second. We conclude, therefore, that the Faraday rotation method is likely to prove the most advantageous satellite method for the study of large-scale ionospheric irregularities, and that it offers important advantages over previous nonsatellite methods.

7. Conclusions

The interpretation of the Faraday rotation of the plane of polarization of a satellite radio signal is at a relatively early stage of development. At the moment, the accuracy of measurements of subsatellite electron content is chiefly limited by uncertainty as to the total number of rotations. In the special case of a spherically stratified ionosphere, or an ionosphere whose gradients are known, this ambiguity can be largely removed. The technique is of considerable value for the study of large-scale irregularities, the interpretation of which is largely independent of ambiguities in the total number of rotations.

It may be expected that the Faraday-rotation methods will be developed still further. In particular, it would be extremely valuable if a satellite specifically designed for the purpose could be launched. Such a satellite should include two transmitters on rather closely spaced frequencies in order to permit resolution of the rotation ambiguity. The satellite should be designed so that the plane of polarization is unaffected by the rotation of the satellite, and some sensing mechanism should be available to define the orientation of the transmitting antenna. With such a satellite, it should be possible to measure the subsatellite electron content to an accuracy of about 1 percent, the accuracy being limited primarily by uncertainties in the magnetic field within the ionosphere. Furthermore, the sensitivity of the method to irregular variations in electron content could be considerably enhanced, particularly in the case of a satellite in a geostationary orbit. For such a satellite, it should be possible to observe variations in the polarization rotation of (say) 0.01 rotation; if this were done at frequency of 20 Mc, it should be possible to measure changes of $\Delta I/I_L$ as small as 1.0×10^{-4}.

The authors are indebted to E. R. Schiffmacher, who was responsible for the radio observations, and to Jane Keltner, who assisted in the analysis of the data. They are also indebted to the Vanguard Computing Center and the Smithsonian Astrophysical Institute for furnishing orbital elements for the satellite, to J. W. Wright of the Central Radio Propagation Laboratory for providing the true-height electron profiles used in our analysis, and to J. W. Warwick of the High Altitude Observatory of the University of Colorado for valuable discussions of the ray-tracing problem. This work was supported by the National Science Foundation under IGY Project No. 32:41, and by the National Aeronautics and Space Administration under Order No. NTF-83. The authors also acknowledge the use of the 20-Mc satellite signals from 1958 Delta 2 (Sputnik III).

8. References

[1] W. A. S. Murray and J. K. Hargreaves, Lunar radio echoes and the Faraday effect in the ionosphere, Nature **173,** 944 (1954).

[2] F. B. Daniels, Electromagnetic propagation studies with a satellite vehicle, ch. 30, Scientific uses of earth satellites, ed. J. A. Van Allen (Univ. of Michigan Press, Ann Arbor, Mich., 1956).

[3] J. A. Ratcliffe, The magneto-ionic theory and its applications to the ionosphere (Cambridge Univ. Press, Cambridge, Mass., 1959).

[4] H. Bremmer, Terrestrial Radio Waves, p. 305 (Elsevier Publishing Co., New York, N.Y., 1949).

[5] R. L. Anderson, Tables of orthogonal polynomial values extended to $N=104$, Agr. Expt. Sta., Iowa State College, Ames, Iowa, Research Bull. 297 (April 1942).

[6] H. F. Finch and B. R. Leaton, The earth's main magnetic field—epoch 1955.0—Monthly Notices, Geophys. Suppl. **7,** 314 (1957).

[7] F. B. Daniels and S. J. Bauer, The ionospheric Faraday effect and its application, J. Franklin Inst. **267,** 187 (1959).

[8] T. G. Hame and E. M. Kennaugh, Recordings of transmissions from the satellite 1958 Δ_2 at the Antenna Laboratory, Ohio State Univ., Proc. IRE **47,** 991 (1959).

[9] R. Parthasarathy and G. C. Reid, Signal strength recordings of the satellite 1958 δ 2 (Sputnik III) at College, Alaska, Proc. IRE **47,** 78 (1959).

[10] G. H. Munro, Travelling disturbances in the ionosphere, Proc. Roy. Soc. **A202,** 208 (1950).

[11] E. N. Bramley and W. Ross, Measurements of the direction of arrival of short radio waves reflected at the ionosphere, Proc. Roy. Soc. **A207,** 251 (1951).

[12] J. D. Whitehead, The focussing of short radio waves reflected from the ionosphere, J. Atmospheric and Terrest. Phys. **9,** 269 (1956).

[13] J. L. Jones, B. Landmark, and C. S. G. K. Getty, Movements of ionospheric irregularities observed simultaneously by different methods. J. Atmospheric and Terrest. Phys. **10,** 296 (1957).

[14] V. V. Vitkevich, Investigation of ionospheric inhomogeneities by radio-astronomical methods, Radiotekhnika i elektronika **3,** 478 (1958) (in Russian).

[15] V. V. Vitkevich and Y. L. Kokurin, Irregular radio-wave refraction and extensive discontinuities in the ionosphere, Radiotekhnika i elektronika **2,** 826 (1957) (in Russian).

(Paper 64D4–6 6

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Note on a Test of the Equivalence Theorem for Sporadic *E* Propagation [1]

J. W. Wright and T. N. Gautier

(February 4, 1960; revised February 23, 1960)

Analysis of two days (123 cases) of sporadic *E* observed simultaneously at oblique and vertical incidence verifies that the classical sec Φ relationship between top frequencies is roughly appropriate for sporadic *E*.

1. Introduction

Medium distance radio propagation (up to 2,000 km) is frequently enhanced through reflection from the sporadic *E* layer (*Es*). At vertical incidence, reflections from *Es* appear to be essentially specular, in the sense that little or no gradual refraction takes place within the layer; this is in contrast to "reflection" from the regular *E* and *F* layers, in which a more gradual refraction process is involved. On the other hand, reflections from *Es* are often observed to be only partial, and the nature of the reflecting surface and, therefore, the reflecting mechanism, are still in some doubt. (See Thomas and Smith [1] [2] for a review of current knowledge of *Es*.) Thus the relationship of vertical-incidence *Es* measurements to oblique-incidence *Es* propagation is also uncertain from a theoretical point of view.

For the simple case of a plane, parallel, earth, and ionosphere, and no magnetic field, the classical "secant law" states that

$$f_{obl}=f_v \sec \Phi,$$

where f_{obl} and f_v are the oblique and vertical incidence frequencies of reflection, respectively, from a plasma frequency $f_p=f_v$, and Φ is the angle of incidence of f_{obl} on the layer.

This relation is applicable not only to radio reflections from a horizontally stratified layer but also to scattering by a stratum of ionospheric irregularities [2]. Thus, there is good reason to expect oblique-incidence *Es* propagation to obey the secant law.

2. Experimental Results

In 1951 and 1952 the National Bureau of Standards operated a pulsed, sweep-frequency, oblique-incidence, ionosphere sounding experiment between Sterling, Virginia, and St. Louis, Missouri [3, 4, 5]. Vertical soundings were obtained simultaneously at the midpoint of this 1,150-km path. One-hop sporadic *E* transmission was frequently observed over this circuit, thereby providing an opportunity to examine the relation between the highest frequencies of reflection observable at oblique and vertical incidence. A preliminary account of the results was included by Smith [6].

Two 24-hr periods were studied for this purpose, providing 123 cases of essentially simultaneous observations of *Es* reflection at vertical and oblique incidence. The ratio of the top frequency observed over the oblique path to that at vertical incidence at the midpoint is shown in figure 1 as a function of *Es* virtual height. The cases in which multiple echoes at vertical incidence had the same highest frequency are indicated on the graphs in order to identify instances of power independence and high reflectivity of the *Es*. The full curve is the secant Φ relationship, plotted versus the height of the midpoint of the triangular path.

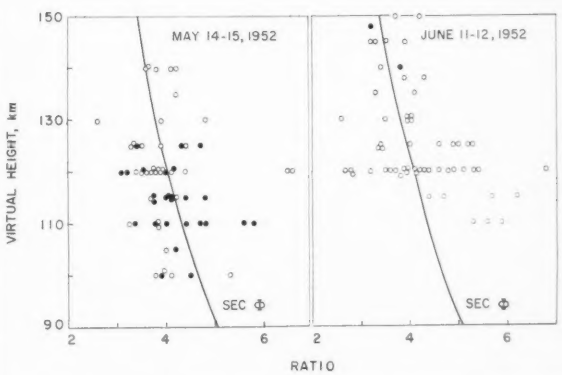


FIGURE 1. *Distribution of ratio of oblique to vertical-incidence sporadic* E *top frequency with virtual height.*

● denotes points for which vertical-incidence multiples extend to same frequency.

[1] Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[2] Figures in brackets indicate the literature references at the end of this paper.

The experimental ratios, f_{obl}/f_v are seen to be in the same range as sec Φ. Whether there is a significant variation of the ratios with height corresponding to sec Φ is a subject for a more detailed analysis. These results, however, lend support to the significance of fEs, the vertical-incidence Es top frequency, as a parameter useful in propagation studies.

3. References

[1] J. A. Thomas and E. K. Smith, A survey of the present knowledge of sporadic *E* ionization, J. Atmospheric and Terrest. Phys. **13,** 295, (1959).

[2] E. V. Appleton and W. J. G. Beynon, An ionospheric attenuation equivalence theorem, J. Atmospheric and Terrest. Phys. **6,** 141 (1955).

[3] P. G. Sulzer and E. E. Ferguson, Sweep-frequency oblique-incidence ionospheric measurements over an 1,150-km path, Proc. IRE **40,** 1124 (1952).

[4] Bernard Wieder, Some results of a sweep-frequency propagation experiment over an 1,150-km east-west path, J. Geophys. Research **60,** 395 (1955).

[5] V. Agy and K. Davies, Ionospheric investigations using the sweep-frequency pulse technique at oblique incidence, J. Research NBS **63D,** 151 (1959).

[6] E. K. Smith, Worldwide occurrence of sporadic *E,* NBS Circ. **582,** 227 (1957).

(Paper 64D4–67)

JOURNAL OF RESEARCH of the National Bureau ot Standards—D. Radio Propagation

Vol. 64D, No. 4, July-August 1960

Daytime Attentuation Rates in the Very Low Frequency Band Using Atmospherics*

W. L. Taylor

(January 25, 1960)

Daytime attenuation characteristics have been computed by comparing the amplitude spectra of atmospheric waveforms recorded at four widely separated stations. The results of these attenuation measurements are presented for the band of frequencies from 3 to 30 kilocycles per second and involving distances of 1,000 to 10,000 kilometers. It was found from these data that attenutation was about 7 to 9 decibels per 1,000 kilometers at 6 kilocycles per second and decreases to about 1 to 3 decibels per 1,000 kilometers at frequencies greater than 10 kilocycles per second. The difference in attenuation rate of west-to-east propagation relative to east-to-west propagation was about 3 decibels per 1,000 kilometers less for frequencies lower than 8 kilocycles per second and about 1 decibel per 1,000 kilometers less from frequencies higher than 10 kilocycles per second.

1. Introduction

Considerable electromagnetic energy is radiated into the atmosphere from lightning discharges. These signals are referred to as atmospherics. The spectra of signals from return stroke lightning discharges attain a peak in the VLF region, usually between 5 and 20 kc. The radiation field spectra near the source are relatively broad i.e., 10 to 15 kc between the frequencies at which the spectral amplitude is 3 db below the peak amplitude. As the propagation distance increases, the spectra begin to narrow such that at distances greater than a few thousand kilometers the spectral width is only 3 to 5 kc. The spectra of atmospherics simultaneously recorded at several locations can be analyzed to determine the attenuation rates as a function of frequency and direction.

Many recent experimental and theoretical investigations have been conducted for purposes of better understanding VLF propagation characteristics.[1] The objective of this paper is to present the results of a study of daytime attenuation rates at VLF using atmospheric waveforms recorded at four widely separated locations.

2. Theoretical Considerations

The Fourier spectrum of an electromagnetic pulse can be represented by the integral

$$E(\omega,d)=\int_0^\infty G(t,d)\cdot e^{-i\omega t}dt \tag{1}$$

where ω is the angular frequency ($2\pi f$); d is the distance in kilometers; t is the time in seconds, and $G(t,d)$ is the instantaneous value of the vertical component of the electric field. Also

$$G(t,d)=0 \text{ for } t<0 \text{ and } t>\tau \tag{2}$$

where τ is the length of the pulse waveform.

Now (1) can be written

$$E(\omega,d)=|E(\omega,d)|e^{-i\phi(\omega,d)}. \tag{3}$$

The function $|E(\omega,d)|$ is the amplitude spectrum and the function $\phi(\omega,d)$ is the phase spectrum as a function of frequency at some distance d. The phase of the spectrum can be used to determine phase velocity of propagation [1].[2] Only the amplitude value, $|E(\omega,d)|$, need be considered for purposes of computing attenuation presented in this paper.

Assuming the source is an equivalent dipole, the spectrum of the vertical electric field at a great-circle distance, d, is given by the mode sum

$$|E(\omega,d)|=A(\omega)W, \tag{4}$$

where $A(\omega)$ is the amplitude coefficient dependent upon the spectrum of the source, and W is the resultant of the complex sum of all active waveguide type modes [2 to 4]. It is extremely difficult to evaluate (4) in a practical situation involving present experimental techniques. However, the main contribution to the spectrum is given by

$$|E(\omega,d)|\simeq A(\omega)[a\sin d/a]^{-1/2}e^{-\alpha(\omega)d}, \tag{5}$$

where a is the radius of the earth in kilometers and $\alpha(\omega)$ is the attenuation coefficient in nepers. This relation expresses the dominant term in a mode series representation of the resultant spectrum, and is an equality only when there is one active mode in the waveguide. This condition was approximately met in the present work as it was in previous work [5].

The principal task was to determine the attenuation as a function of frequency. To accomplish this the vertical electric field waveform of the same lightning stroke was recorded at several locations. The Fourier spectrum of each waveform was then

* Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[1] A selection of representative references is appended.

[2] Figures in brackets indicate the literature references at the end of this paper.

calculated. Using the subscript i to denote the far station and the subscript j to denote the near station, such that $d_i > d_j$, it follows from eq (5) that

$$\exp\left[\alpha(\omega)\cdot(d_i-d_j)\right]=\frac{|E(\omega,d_j)|}{|E(\omega,d_i)|}\cdot\left[\frac{\sin(d_i/a)}{\sin(d_j/a)}\right]^{-1/2}. \quad (6)$$

The attenuation coefficient $\alpha'(\omega)$ expressed in decibels per 1,000 km is given by

$$\alpha'(\omega)=\left[20\log\frac{|E(\omega,d_j)|}{|E(\omega,d_i)|}-10\log\frac{\sin d_i/a}{\sin d_j/a}\right]\cdot\frac{10^3}{d_i-d_j} \quad (7)$$

where $\alpha'(\omega)=\alpha(\omega)\times 8.68\times 10^3$.

3. Experimental Procedure

3.1. Equipment

The vertical component of the electric field of atmospherics was received using a vertical antenna and wideband receiving equipment. These atmospheric waveforms were displayed on a multigun cathode-ray tube at two sensitivities and recorded on continuously moving photographic film. The amplitude response of the waveform channel was approximately flat throughout the band-pass and sloped to 6-db points at frequencies of 1 and 100 kc. The phase response was essentially linear with frequency within the band-pass. The cathode-ray tube display is proportional to the free-space field incident at the antenna. Included in the vertical electric field channel was a 23-ft vertical antenna, antenna cathode follower, band-pass filters composed of RC networks, a 24-μsec delay line, wideband amplifier, and oscilloscope. The overall characteristics of this channel were adequate to faithfully record the VLF components of atmospheric waveforms.

Calibration was performed by inserting a voltage into a dummy antenna driving the cathode follower, and adjusting amplifier gain to give the desired deflection on the cathode-ray tube. This calibration voltage was equal to the expected field strength times the antenna effective height, assumed to be one-half its physical length.

Amplitude threshold triggering was used to discriminate against the numerous low-amplitude atmospherics. Signals that exceeded a selected amplitude, prior to passing through the delay line, activated the sweep circuit of the oscilloscope. All signals received with the vertical antenna were displayed on the oscilloscope from 24 μsec before to 1 msec after the time when the amplitude of the atmospheric reached the triggering level. Use of the delay line minimized possible loss of the early part of the waveforms.

The direction of arrival of atmospherics was displayed on the same multigun oscilloscope. A pair of vertical electrostatically-shielded loop antennas, arranged at right angles to each other, loop receivers, filters, delay lines, and amplifiers composed the direction-finding system. The overall transfer characteristics of each loop-antenna channel closely approximated that of the vertical-antenna channel. The differences between the amplitude and phase responses of the two loop-antenna waveform channels were less than 0.5 db and 6 deg, respectively, over the entire band-pass.

The direction-finding display on the oscilloscope was activated for approximately 200 μsec at the beginning of each atmospheric. Thus, the direction of arrival was determined from the initial portion of the signal, which is essentially vertically polarized. Accuracy of determining direction of arrival was about $\pm 2°$ for each station.

The time of arrival of atmospherics was determined to and accuracy of about ± 1 msec. This was accomplished by recording on the waveform records the standard time signals from stations WWV (or WWVH) and timing marks from a secondary frequency standard which was synchronized with the seconds "tick" from WWV. The measured arrival times of atmospherics were corrected for propagation time of the signals from the standard time station to the recording site.

Four sets of this equipment were operated simultaneously at Boulder, Colo.; Salt Lake City, Utah; Palo Alto, Calif.; and Maui, Hawaii, on a number of occasions during the summer of 1958. The recording sensitivities and the triggering amplitudes were selected for each station to accommodate atmospherics originating in the western half of the United States. The transfer characteristics of the vertical antenna waveform channel were frequently checked at each station to insure uniformity of responses between stations. It was found that the differences between the response characteristics of the equipment at different stations were less than ± 0.5 db and $\pm 20°$, respectively, in amplitude and phase. These differences are small for the purposes of this paper.

Atmospheric waveforms resulting from a particular lightning discharge were located on the photographic records made of each station by virtue of their times of reception and directions of arrival. The locations of these discharges were determined by triangulation, using the direction of arrival indicated at each station.

3.2. Analysis—East-to-West Propagation

Waveforms from fifty lightning discharges were obtained from the data collected on July 2, 1958, 1900 to 2100 G.m.t.; July 7, 1958, 2300 to 0100 (July 8) G.m.t.; and July 9, 1958, 1900 to 2100 G.m.t. Atmospherics recorded at each of the four stations from twenty of these lightning discharges were analyzed. This analysis consisted of (1) scaling the individual waveforms recorded at each station to determine the functions $G(t, d)$, (see eqs (1) and (2)), (2) computing the complex spectra (eq (1)) every kilocycles per second from 1 to 100 kc, and (3) computing the attenuation (see eq (7)) for each pair of stations, i.e., Boulder—Maui, Salt Lake City—Maui, and Palo Alto—Maui. The distances from the sources of each atmospheric to the near station were between

1,000 to 2,400 km, while the distances from the sources to Maui (the far station) were between 4,160 to 5,200 km. Accuracy of determining the location of the individual atmospherics was about ± 10 percent of the distance to the near station.

One set of waveforms, their corresponding amplitude spectrum, and the computed attenuations ($\alpha'(\omega)$), representative of those used in this analysis, are given in the first four figures. The waveforms,

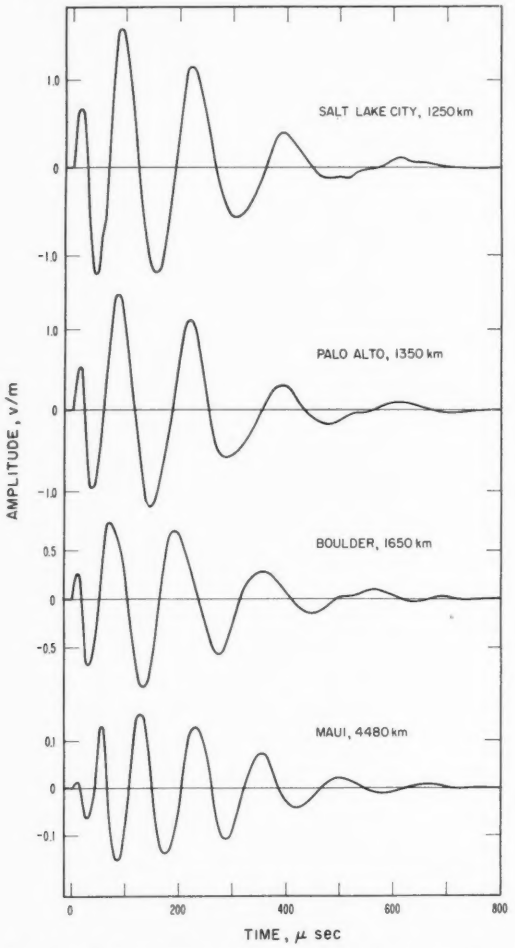


FIGURE 1. *Representative atmospheric waveforms, east to west.*

shown in figure 1, are reproduced on the same time base but different amplitude scales, while the recording stations and the distances to the source are indicated. Note in particular the progression to shorter time intervals of the points where each waveform successively crosses the zero amplitude axis as propagation distance increases. The times in microseconds of these quasi half-periods, presented in figure 2, show clearly the progression to shorter time intervals as propagation distance increases. These values, from the representatives set of waveforms, agree quite reasonably with those found by other

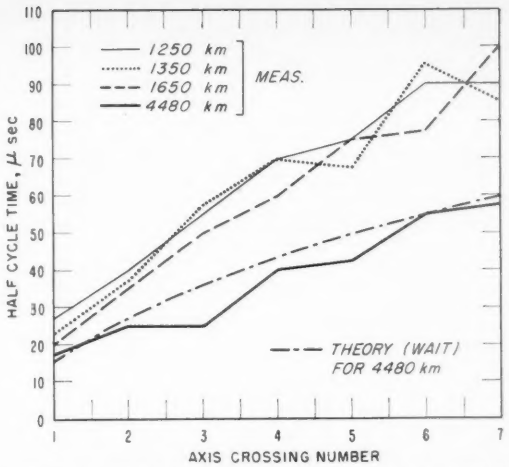


FIGURE 2. *Quasi half-periods of representative atmospherics, east to west.*

workers [6, 7], and the behavior is in accord with theoretical investigation of the phenomenon [8]. Further elaboration concerning the atmospheric wave shapes are beyond the scope of this paper.

The spectra of these atmospherics are presented in figure 3. It should be observed that the frequency of

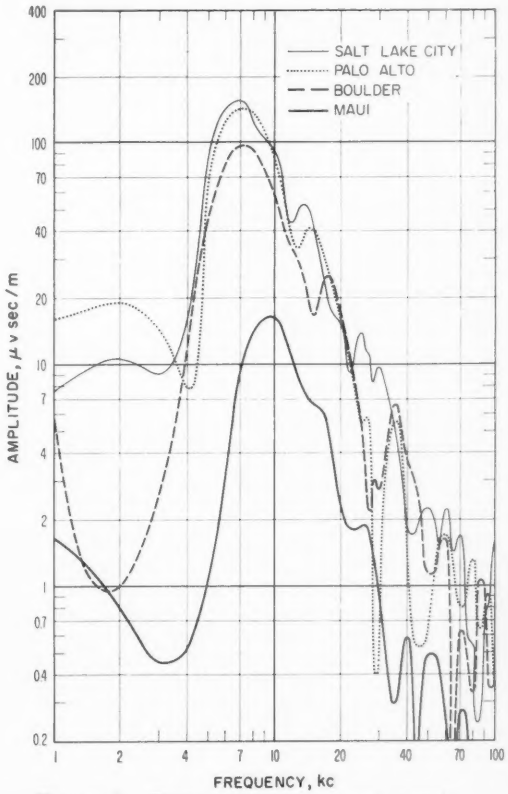


FIGURE 3. *Representative spectra, east to west.*

the peak amplitude of each spectrum is increasing as the propagation distance increases.

The attenuation in decibel per 1,000 km computed from these spectral values are presented in figure 4. Values of attenuation for each frequency from each set of computations, using Boulder to Maui, Salt Lake City to Maui, and Palo Alto to Maui, are in reasonably good agreement. Accuracy of determining differences in distance between each station of a pair and the source was about 7 percent.

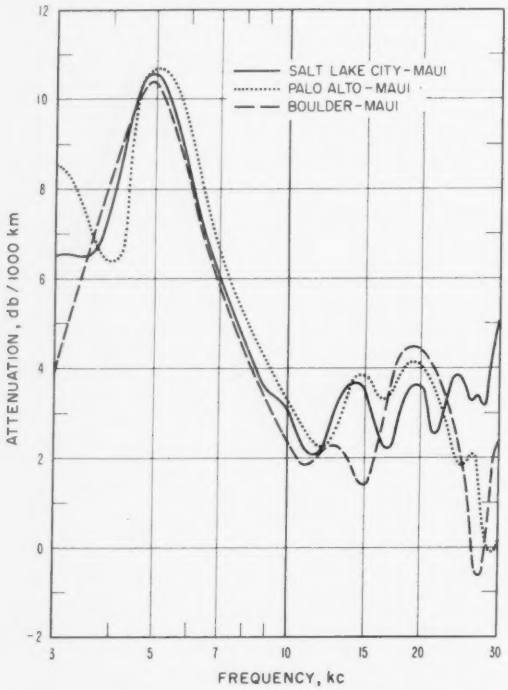


FIGURE 4. *Representative attenuation, east to west.*

Values of the spectral components of frequencies lower than about 4 kc and higher than about 30 kc are relatively small, less than 20 db below the peak spectral value. Errors in the computed attenuation resulting from limitations in the described technique may become large as the amplitude of the spectra approaches very small values. Therefore, computations of attenuation were limited to the VLF region of the spectrum (3 to 30 kc).

The analysis of the waveforms from the other lightning discharges give results similar to the example presented in figures 1, 3, and 4.

Considerable variability was found between the shapes of the attenuation versus frequency curves, obtained from the data recorded at each pair of stations. Six individual curves of attenuation as a function of frequency are presented in figure 5. The curves are identified by capital letters (A) and their corresponding "zero decibel level" by lower case letters (a). The names of the stations and distances from the source in kilometers are given with the

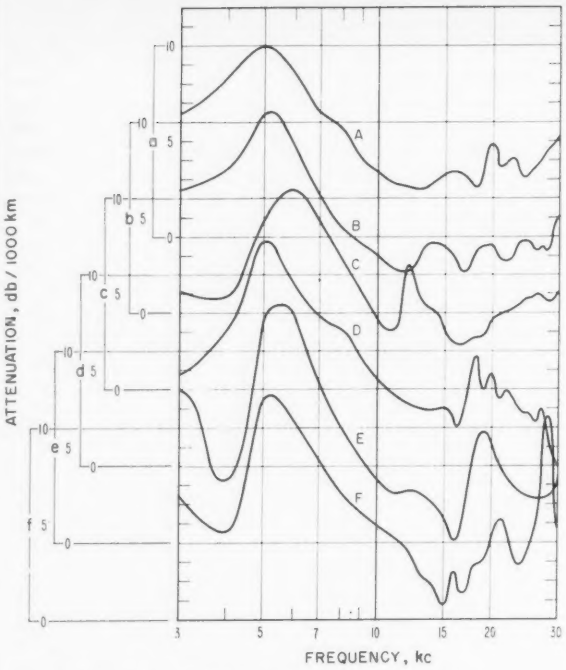


FIGURE 5. *Individual attenuation, east to west.*

A. Palo Alto, 1,040 km to Maui, 4,160 km.
B. Salt Lake City, 1,250 km to Maui, 4,480 km.
C. Salt Lake City, 1,470 km to Maui, 4,700 km.
D. Boulder, 1,650 km to Maui, 4,300 km.
E. Palo Alto, 1,840 km to Maui, 4,330 km.
F. Palo Alto, 2,000 km to Maui, 4,320 km.

figure. Variations of the computed attenuation about some mean value can be seen between 10 to 30 kc for each of these curves. The small, smooth variations of less than ± 1 db at frequencies above 10 kc, observed in curve B may well be produced by interference from higher order modes. Abrupt changes in attenuation, below 5 kc in all curves, and observed near 12 kc in curve C, 17 kc in curve D and E, and 29 kc in curve F, are most likely produced by various limitations in the recording and analysis techniques. These abrupt changes were observed only at those frequencies where the spectral amplitude component was extremely small for one or both of the waveforms recorded at a pair of stations. Small unwanted signals and noise, superimposed on the waveforms, and errors in tracing and scaling the waveforms may vary the resultant amplitude spectra by a few microvolt-seconds per meter (μvsec/m) at any frequency. Therefore, large errors would be expected at those frequencies where the computed amplitude of the spectra were small and therefore approaching the value introduced through limitations in the techniques.

Errors in the computed attenuation will tend toward a random distribution. Averaging the attenuation computed from a relatively large sample of independent data should give attenuation values that are reasonably accurate.

The average attenuation-versus-frequency values computed from the data recorded at each pair of stations are presented in figure 6. The average location of all lightning discharges used in this study was at 51°N to 118°W, i.e., the center of gravity for the discharges. The average distance from the lightning discharges was 1,220 km to Salt Lake City, 1,470 km to Palo Alto, 1,530 km to Boulder. All of these propagation paths occurred over land. The average distance to Maui was 4,580 km, which was all sea water except for 480 km at the source end of the average path. It can be seen from these distances that the average propagation path, over which these attenuation values were computed, was about 82 percent sea water for the Salt Lake City to Maui path, 76 percent sea water for the Palo Alto to Maui path, and 74 percent sea water for the Boulder to Maui path. The midpoint between this center or average location and Maui was approximately 34° N to 139° W where the direction of propagation was about 230° from geographic north.[3]

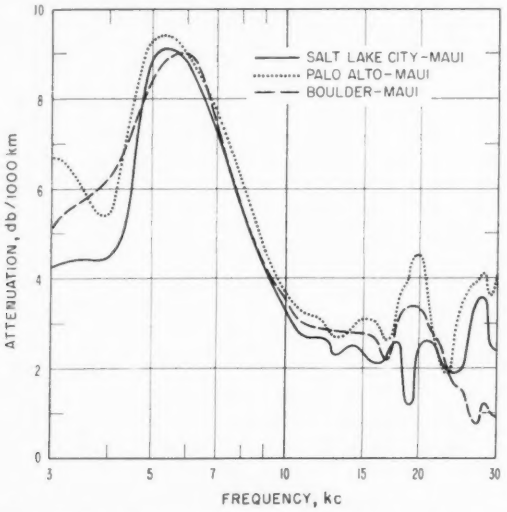


FIGURE 6. *Average attenuation by station, east to west.*

The average of the three curves is shown in figure 7. This curve represents experimentally-derived attenuation coefficients versus frequency for an approximately east-to-west path over predominantly sea-water conditions for summer, daytime conditions. The attenuation values represented by this curve may be low compared to an all sea-water path. An approximate correction to these values can be calculated by assuming the attenuation for a land path is 1 db greater for all frequencies than for a sea-water path. Using this approach, it can be shown that the attenuation values should be possibly increased by about 0.3 db/1,000 km.

3.3. Analysis—West-to-East Propagation

Attenuation coefficients for this study were computed from the data recorded at Maui and Palo Alto.

[3] Geomagnetic direction of propagation was about 215° and the dip angle of the earth's magnetic field was about 56°.

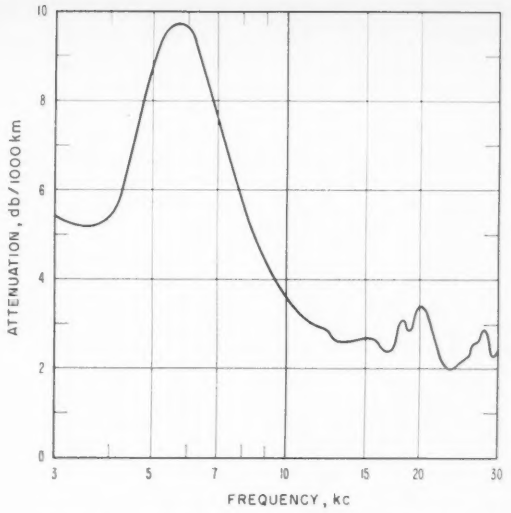


FIGURE 7. *Average daytime attenuation, east to west.*

Only one set of attenuation values were obtained from each lightning discharge, in which Maui was the close station and Palo Alto was the far station. The stations at Salt Lake City and Boulder did not record the atmospherics used in this part of the analysis because of the high level of atmospheric activity relatively near these stations during each observation period. Waveforms from 16 lightning discharges occuring in the western Pacific Ocean were obtained from the data collected on July 2, 1958, 1900 to 2100 G.m.t., and on July 9, 1958, 1900 to 2100 G.m.t. Eight of these were selected for the attenuation study.

Analysis of these data was conducted in the same manner as for the east-to-west attenuation computations. One pair of representative waveforms is shown in figure 8, reproduced on the same time base. Again the times of the zero-axis crossings of the Palo Alto waveform, recorded at the more distant station, are slightly less than for the Maui waveform.

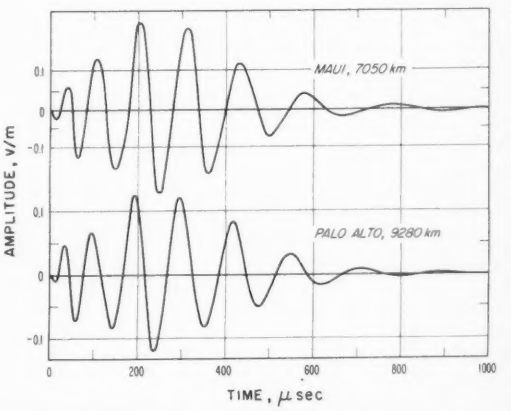


FIGURE 8. *Representative atmospheric waveforms, west to east.*

The spectra of these two atmospherics are presented in figure 9. The frequency of the peak amplitude of the Palo Alto spectrum occurs about 500 cycles higher than the corresponding peak of the Maui spectrum. Oscillations of the spectra, occurring between about 15 to 40 kc, are probably produced, at least in part, by an apparant "phase" shift in the waveform between the fifth and sixth half cycles as seen in figure 8. Phenomena of this type have been reported in the literature [9].

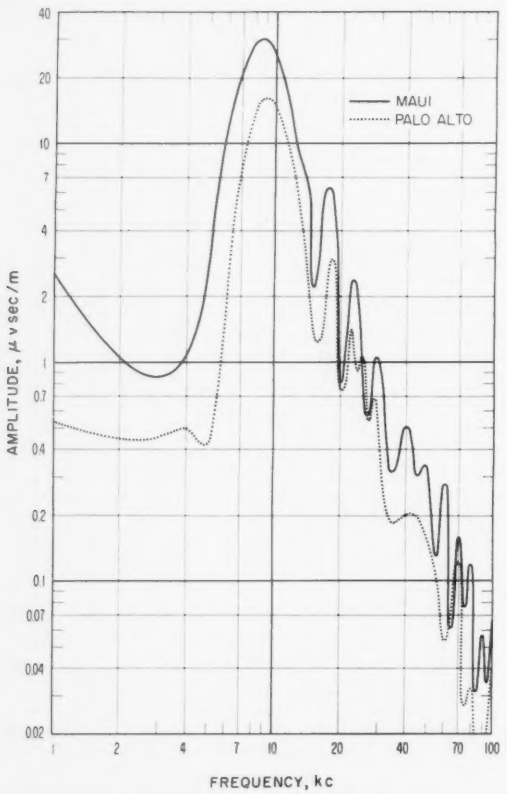


FIGURE 9. *Representative spectra, west to east.*

Attenuation in decibels per 1,000 km computed from this pair of atmospherics is shown in figure 10. Oscillations in the Palo Alto and Maui spectra, mentioned above, show their effect here as oscillations in the attenuation curve.

The average attenuation values, computed from the selected pairs of waveforms, are presented in figure 11. The accuracy of determining the location of individual atmospherics was about ±20 percent of the distance to the near station (Maui), but the accuracy of determining the differences in distance between each station and the source was about ±10 percent. Approximate average distance from the source to Maui was 6,460 km, and from the source to Palo Alto was 9,160 km. The midpoint between the center or average location of the discharges and

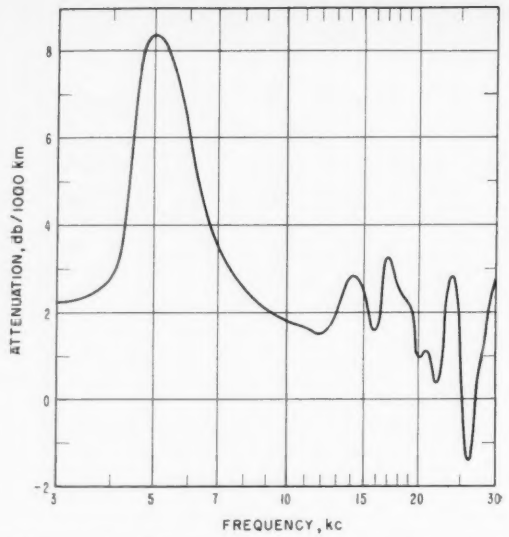


FIGURE 10. *Representative attenuation, west to east.*

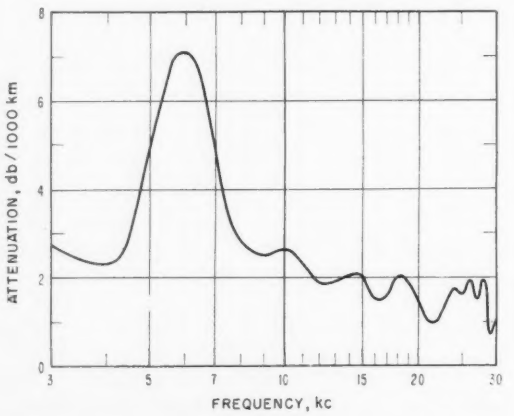


FIGURE 11. *Average daytime attenuation, west to east.*

Palo Alto was approximately 33° N to 176° W where the direction of propagation was about 68° from geographic north.[4] This curve represents experimentally derived attenuation coefficients versus frequency for an approximately west-to-east path, over sea water, for summer, daytime propagation.

4. Concluding Remarks

The general shape and the absolute values of the east-to-west attenuation-versus-frequency curve in figure 7 agrees reasonably well with the experimental and theoretical works of others.[5] The attenuation values for west-to-east propagation in figure 11 are somewhat lower than might be expected for a sharply bounded ionospheric model having reasonable values

[4] Geomagnetic direction of propagation was about 58° and the dip angle of the earth's magnetic field was about 49°.
[5] A selection of representative references is appended.

of reflection height, electron density, and collisional frequency. The differences between these two attenuation curves are shown in figure 12. This represents the attenuation for east-to-west propagation minus the attenuation for west-to-east propagation. A signal propagating from west to east is attenuated 3 db/1,000 km less between frequencies of about 3 to 8 kc, and 1 db/1,000 km less between frequencies of about 10 to 30 kc, than a signal propagating from east to west. The effect of a decreased attenuation at frequencies below about 8 kc would be to enhance the smooth oscillatory characteristics of atmospherics at great ranges. This characteristic has been observed by Hepburn [6] and by Chapman and Pierce [10] who reported the occurrence of long oscillatory, quasi-sinusoidal waveforms from a southwesterly direction.

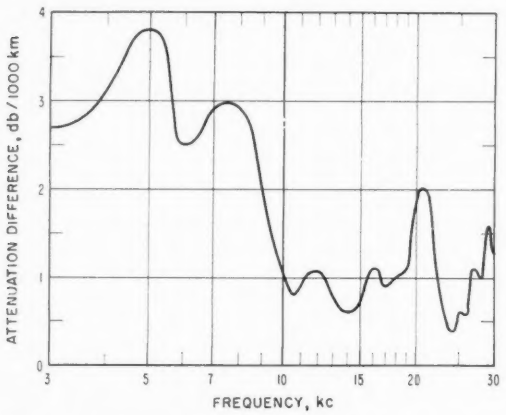


FIGURE 12. *Increased attenuation for east to west compared with west to east for daytime propagation.*

The 1 db/1,000 km difference in attenuation rates between 10 and 30 kc is also in agreement with various propagation studies using transmissions from VLF stations. Attenuation rates [11] computed from the data of Round et al. [12], indicate that propagation from east to west is characterized by an increase in attenuation of 1 or 2 db/1,000 km compared with propagation from west to east. Measurements of Heritage et al. [13], indicates that attenuation over the Pacific is slightly higher for east-to-west than for west-to-east propagation. Similar results were reported by Crombie [14].

It seems reasonable to conclude that propagation from west to east is better than from east to west. Much work remains to be done if this effect is to be understood fully. In particular, attenuation versus frequency should be determined as a function of the direction of propagation with respect to the earth's magnetic field lines. Theoretical investigations also should be extended to include those effects.

The author thanks C. Alley, University of Utah at Salt Lake City; W. Cornell, Stanford University at Palo Alto; C. Chilton, NBS at Boulder; S. Katahara, NBS at Maui; and all others who assisted in the collection of data for this paper. Also, the author thanks A. G. Jean and J. R. Wait for stimulating discussions on this subject, H. H. Howe who devised the program for the spectra computations, L. J. Lange and A. Murphy for further assistance with the computations, and Lois Pollock for computing attenuation values and preparing the illustrations.

5. References

[1] A. G. Jean, W. L. Taylor, and J. R. Wait, VLF phase characteristics deduced from atmospheric waveforms, J. Geophys. Research **65,** 907 (March 1960).
[2] J. R. Wait, The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760 (1957).
[3] J. R. Wait, On the mode theory of VLF ionospheric propagation, Geofis. pura e appl. (Milano) **37,** 103 (1957).
[4] J. R. Wait, An extension to the mode theory of VLF ionospheric propagation, J. Geophys. Research **63,** 125 (1958).
[5] W. L. Taylor and L. J. Lange, Some characteristics of VLF propagation using atmospheric wave forms, Recent Adv. Atmos. Elec., p. 609 (Pergamon Press, Inc., New York, N.Y., (1959)).
[6] F. Hepburn, Wave-guide interpretation of atmospheric waveforms, J. Atmospheric and Terrest. Phys. **10,** 121 (1957).
[7] H. Norinder, The waveform of the electric field in atmospherics, Arkiv Geophys. **2,** 161 (1954).
[8] J. R. Wait, Propagation of VLF pulses to great distances, J. Research NBS **61,** 187 (1958) RP2898.
[9] F. Hepburn, Interpretation of smooth type atmospheric waveforms, J. Atmospheric and Terrest. Phys. **14,** 262 (1959).
[10] J. Chapman and E. T. Pierce, Relations between the character of atmospherics and their place of origin, Proc. IRE **45,** 804 (1957).
[11] J. R. Wait, A study of VLF field strength data, Geofis. pura e appl. **41,** 73 (1958).
[12] H. J. Round, T. L. Eckersley, K. Tremellen, and F. C. Lunnen, Report on measurements made on signal strength at great distances during 1922 and 1923 by an expedition sent to Australia, J. Inst. Elect. Engrs. **63,** 933 (1925).
[13] J. L. Heritage, S. Weisbrod, J. E. Bickel, A study of signal-vs-distance data at VLF Record Symp. on VLF propagation, Boulder, Colo. (Jan. 1957).
[14] D. D. Crombie, Differences between east-west and west-east propagation of VLF signals over long distances, J. Atmospheric and Terrest. Phys. **12,** 110 (1957).

Additional References

P. W. A. Bowe, The waveforms of atmospherics and the propagation of VLF radio waves, Phil. Mag. 42, 121 (1951).
K. G. Budden, On the propagation of a radio atmospheric, II, Phil. Mag. **43,** 1179 (1952).
F. W. Chapman and R. C. V. Macario, Propagation of audiofrequency radio waves to great distances, Nature **177,** 930 (1956).
W. L. Taylor and A. G. Jean, Very-low-frequency radiation spectra of lightning discharges, J. Research NBS **D63,** 199 (1959).
J. R. Wait, The attenuation vs frequency characteristics of VLF radio waves, Proc. IRE **45,** 768 (1957).
A. D. Watt and E. L. Maxwell, Characteristics of atmospheric noise from 1 to 100 kc, Proc. IRE **45,** 787 (1957).

(Paper 64D4–68)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Measured Electrical Properties of Snow and Glacial Ice [1,2]

A. D. Watt and E. L. Maxwell

(January 22, 1960)

The electrical properties of snow and glacial ice near 0° C have been observed over the frequency range from 20 cycles per second to 200 kilocycles per second. In general, the conductivity of snow and glacial ice is found to be much higher than that for pure ice. This is particularly so at frequencies below 2 kilocycles per second.

The magnitude of the complex conductivity for glacial ice appears to increase with temperature at frequencies below 200 cycles per second and to decrease with temperature above this frequency.

1. Introduction

The electrical properties (i.e., conductivity and permittivity) of snow and ice can have an appreciable effect upon the propagation of electromagnetic waves over such media. Similarly, the characteristics of ground-based antennas may be severely modified [1, 2].[3] Although the properties of pure ice as measured in the laboratory are relatively well known, the physical characteristics and formation of snow and glacial ice are different, and thus it is important that their properties be measured under actual field conditions. Considerable dispersion exists in the electrical properties of ice throughout the audiofrequency range, and in order to permit measurements of the electrical properties of snow and glacial ice in their natural state, two independent methods have been employed over the frequency range from 20 cps to 200 kc. These are described in the following section and the resulting electrical properties presented.

2. Measurement Methods and Nomenclature

The bridge and substitution method shown in figure 1 permits observations to be made on the conductivity and permittivity (dielectric constant) of a lossy dielectric. Numerous forms of electrodes have been employed including parallel plates and the parallel rods shown. The parallel rods have the great advantage of permitting their insertion into the medium by the simple expedient of drilling two holes with an ice auger. This method produces a minimum amount of disturbance in the medium, and when the experiment is carefully conducted the effects of fringe capacity etc., are greatly minimized. This method suffers from one primary deficiency in that the electrodes must fit very closely in the ice, and there is always the question of possible errors in

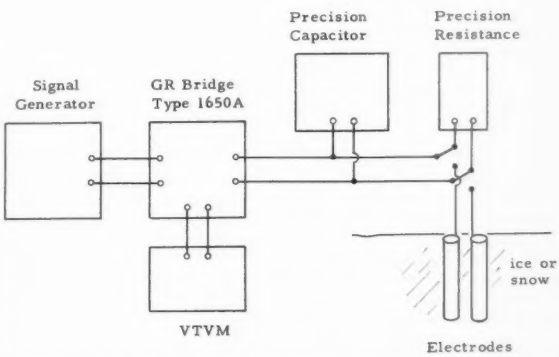


FIGURE 1. *Bridge and substitution method for measuring conductivity and dielectric constant.*

the determination of conductivity and permittivity due to interface or surface-layer effects on the electrodes themselves.

The other method employs two pairs of electrodes. It is described generally by Wait [3] and has been used in measuring ground conductivity in the audiofrequency region. The use of four electrodes permits a determination of the magnitude of the complex conductivity $|\sigma|$ to be made in such a manner that surface-layer effects [4] at the electrodes are eliminated. At least three basic arrangements of electrodes were employed during these measurements including the Wenner, Eltran, and right-angle arrays. The Eltran arrangement shown in figure 2 proved to be the convenient form to use since it permits a ready changing of electrode spacings, and accurate results can be obtained over a relatively wide range of electrode spacings provided that the precaution of elevating the wire several feet above the surface of the snow or ice is taken. If the wires connecting the electrodes were not elevated, the large dielectric constant of the snow and the rather high terminal resistance resulted in a nonuniform distribution of current along the wire. This, of course, invalidates the relations employed in calculating $|\sigma|$ based upon a model where the current and potential electrodes

[1] Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

[2] The studies contained in this paper were sponsored by the Office of Naval Research under Contract No. NR 371-291.

[3] Figures in brackets indicate the literature references at the end of this paper.

are discrete points. At very short spacings, the results with all three methods were similar.

It is important that the vacuum-tube voltmeter across the potential electrodes have a very high input impedance and that the voltmeter stake impedances be made as low as possible by pouring salt or salt water around electrodes placed in the ice so that the voltage indicated is equal to the true potential difference between the potential electrodes. These precautions are necessary since the stake impedances are effectively in series with the potential source whose voltage is desired. In some cases, it may be difficult to reduce the stake impedances below 0.5 meg and corrections of the voltmeter readings may be necessary [4].

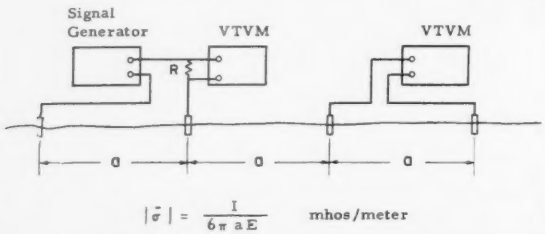


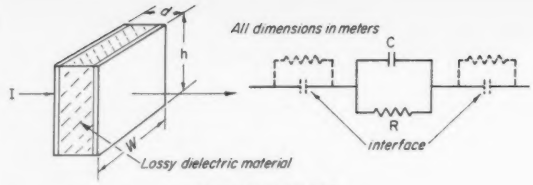
FIGURE 2. *Four electrode [ELTRAN] array for measuring the magnitude of the complex conductivity.*

The signal generator should be capable of producing relatively high voltages so that sufficient current can be induced into the current electrodes. This makes the voltage produced at the potential electrodes appreciably greater than the background noise caused by atmospheric noise and earth-current fields. Signal generators with outputs of 1 v have been employed for spacings up to 20 ft while beyond this it has been found necessary to employ an additional power amplifier with a capability of supplying 300 v into 6,000 ohms. This arrangement has permitted electrode spacings of up to 1,000 ft provided that some filtering is employed at the vacuum-tube voltmeter to eliminate the fields produced by VLF transmitting stations.

The four-electrode method using the instrumentation shown eliminates the effects of surface phenomena at the electrodes but does not give the magnitude of the conductivity and permittivity independently.[4] The magnitude of the complex conductivity employed by this means can, however, be compared with that obtained by the bridge substitution method as a check to see whether or not surface phenomena at the electrodes in the bridge measurements was great enough to cause an error in the observations.

The physical characteristics of the electrodes, used in the bridge method, and the relation between the equivalent circuit and the basic properties of the dielectric material are shown in figure 3. The two basic physical properties to be measured are the conductivity σ expressed in (mhos/meter), and the permittivity (dielectric constant), ϵ, expressed in

[4] This can be accomplished if the relative phase between current I and voltage V is measured [4].



LIST OF SYMBOLS

σ	=	Conductivity (mhos/meter)
ϵ	=	Permittivity, i.e., dielectric constant (farads/meter)
A^*	=	ωh, plate area (meters^2)
R	=	$d/A^*\sigma = \frac{V}{I}$, resistance (ohms)
G	=	$1/R = \sigma A/d$, conductance (mhos)
C	=	$\epsilon A/d$, capacitance (farads)
A	=	Effective area $\propto A^*$ when $d \ll \omega$ and h
X_c	=	$-\frac{1}{\omega c}$, reactance (ohms)
B	=	$-\frac{1}{X} = \omega c$, susceptance (mhos)
Y	=	$G + jB$, admittance (mhos)
	=	$\sigma A/D + j\omega\epsilon A/D$
$\bar{\sigma}$	=	$\sigma + j\omega\epsilon$ complex conductivity "admittivity" (mhos/meter)
$\lvert\bar{\sigma}\rvert$	=	$\sqrt{\sigma^2 + (\omega\epsilon)^2}$ magnitude of complex conductivity "admittivity" (mhos/meter)
$\bar{\epsilon}$	=	$\epsilon' - j\epsilon''$ complex dielectric constant
ϵ'	=	$\epsilon/\epsilon_o = C/C_o$ real part of relative dielectric constant
ϵ''	=	$\sigma/\omega\epsilon_o = \frac{1}{R\omega C_o}$ imaginary part of relative dielectric constant
σ	=	$5.56 \times 10^{-11} \epsilon'' \times f(c/s)$ (mhos/meter)

FIGURE 3. *Lossy dielectric capacitor and equivalent circuit.*

(farads/meter). When the plates or rods are immersed in an effectively infinite medium, it is necessary to use the effective area to separation ratio A/d rather than the ratio of physical values A^*/d shown in the illustration.

It is evident that

$$\sigma = \frac{1}{R\,A/d} \tag{1}$$

and that

$$\epsilon = \frac{C}{A/d} \tag{2}$$

where σ is the conductivity in mhos/meter, R is the equivalent parallel resistance in ohms, A is the effective area in square meters, d is the effective separation in meters, ϵ is the permittivity of the substance in farads/meter, and C is the capacity in farads. Since the ratio A/d may influence the interface effect, the value of this ratio is given for each electrode arrangement employed. Although A/d can be calculated with fairly good accuracy, the value determined experimentally from the air capacity of the electrodes is used in all the calculations of this paper. Most of the nomenclature employed is now well standardized [4]. The complex conductivity, $\bar{\sigma}$, may not be well known and the possibility of employing "admittivity" in future work is suggested. It should be emphasized that the quantity obtained by the four-electrode method is $|\bar{\sigma}|$ the magnitude of

the "admittivity" rather than the conductivity σ which is the dominant term in most ground measurements.

3. Properties of Pure Ice

The electrical and physical properties of ice have been discussed in considerable detail by Dorsey [5]. He presents tabulations of permittivity and conductivity as reported by numerous observers. A more recent paper by Auty and Cole [6] has analyzed the lack of close agreement between measured dielectric constants of ice and has found that in general they appear to be attributable to the formation of voids or cracks in the samples being measured.

For a material with a single relaxation frequency the simple theory of dielectric relaxation predicts a complex dielectric constant $\bar{\epsilon}$

$$\bar{\epsilon}=\epsilon'-i\epsilon''=\epsilon_\infty+(\epsilon_{dc}-\epsilon_\infty)/(1+if/f_c) \qquad (3)$$

where the time factor is $e^{i\omega t}$, ϵ_{dc}, and ϵ_∞ are the so-called equilibrium and limiting high-frequency values of relative dielectric constant, f is the frequency at which the electrical properties are being measured, and f_c is defined as the relaxation frequency which is equal to $1/\tau$, where τ is the relaxation time. The complex plane presentation of $\bar{\epsilon}$ forms semicircles as shown in figure 4. This type of presentation is known as the Argand diagram. Auty and Cole [6]

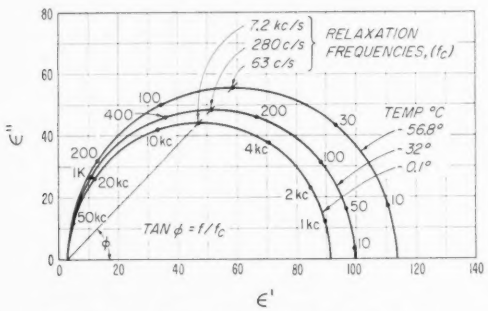


FIGURE 4. *Electrical properties of pure ice.*
From Auty and Cole.

with carefully designed equipment in experimental procedures found that pure ice does in fact adhere very closely to the form predicted by this simple theory, and they were able to obtain results with a high degree of repeatability. From their data we have obtained the relaxation frequency f_c as a function of temperature as shown in figure 5. The high-frequency dielectric constant ϵ_∞ was found to be very close to 3 for all temperatures ranging from $-0.1°$ to $-65.8°$ C; and ϵ_{dc}, also shown in figure 5, has values ranging from approximately 90 to 114 over the temperature range indicated. The value of ϵ corresponding to the relaxation frequency, ϵ_c, permits a rapid construction of Argand diagrams since this point represents the center of the circular diagram. It is obvious that the frequency at the

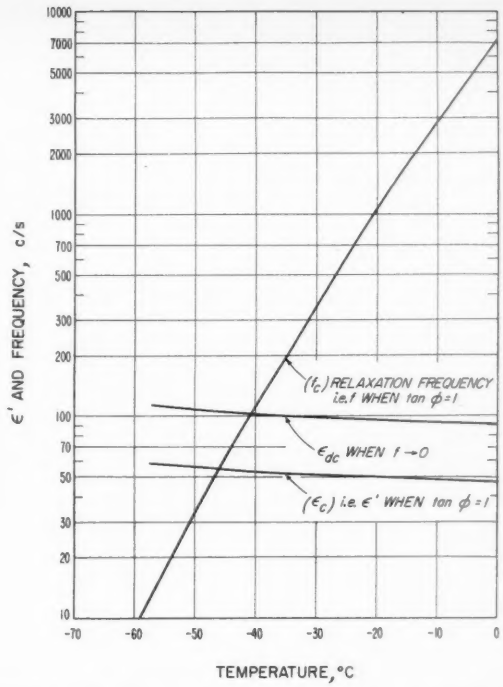

ϵ' AND FREQUENCY, c/s
(f_c) RELAXATION FREQUENCY i.e. f WHEN tan ϕ=1
ϵ_{dc} WHEN f→0
(ϵ_c) i.e. ϵ' WHEN tan ϕ=1
TEMPERATURE, °C

FIGURE 5. *Pure ice—frequency and relative permittivity when tan ϕ=1 versus temperature.*
From Auty and Cole.

top of the semicircular diagram will be the relaxation frequency. Frequencies at other points on this diagram can readily be obtained by the relationship

$$\tan \phi = f/f_c \qquad (4)$$

where ϕ is the angle indicated in figure 4. The relative dielectric constant $\epsilon'=\epsilon/\epsilon_0$ can be readily obtained from diagrams such as figure 4 and the results presented as a function of frequency for various temperatures as shown in figure 6.

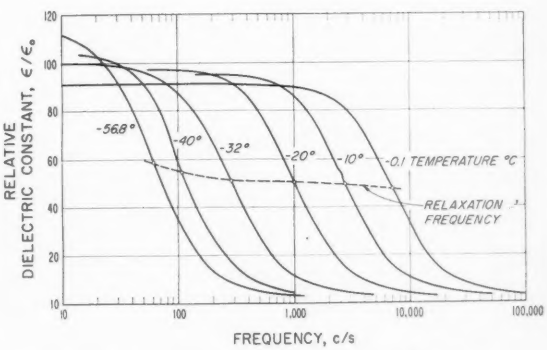


FIGURE 6. *Relative dielectric constant of pure ice.*
Calculation based on Auty and Cole.

It is interesting to note that the equilibrium relative dielectric constant ϵ_{dc} increases with decrease in temperature. This same general increase in ϵ_{dc} has been predicted theoretically by Powles [7]. The actual magnitude, however, is either below or somewhat above the values obtained experimentally depending upon the type of theoretical model employed in the calculations.

The conductivity of pure ice can be obtained from the Argand diagram of figure 4 by means of the relationship

$$\sigma = \epsilon'' \omega \epsilon_0 \tag{5}$$

which can also be written as

$$\sigma = 5.56 \times 10^{-11} \epsilon'' f \tag{6}$$

where f is the frequency in cycles per second and σ is the conductivity in mhos/meter. The results of these calculations, shown in figure 7, indicate that for a given temperature and at frequencies below the relaxation frequency that the conductivity is directly proportional to the square of the frequency.

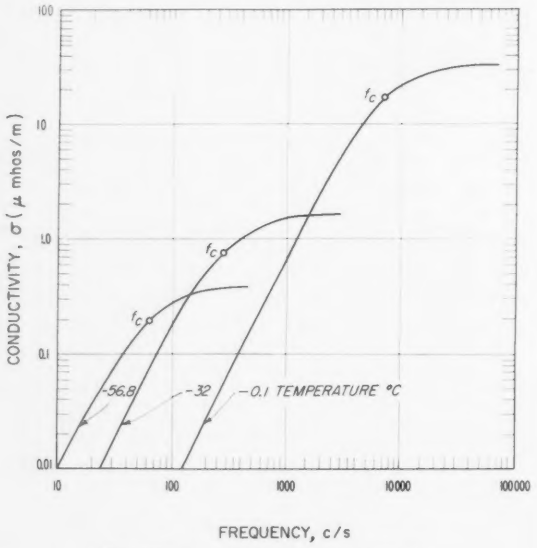


FIGURE 7. *Conductivity of pure ice.*
Calculations based on Auty and Cole.

4. Experimental Results

4.1. Snow Drifts in Colorado

The properties of drifted snow were measured at two locations in the mountains of Colorado. The Argand diagrams resulting are shown in figure 8 where it is obvious that the characteristics are different from those of pure ice. Four electrode methods of measuring the absolute magnitude of the complex conductivity are also employed and a comparison of results is shown in figure 9, where it is seen that relatively good agreement exists between the two types of measurements. The type of

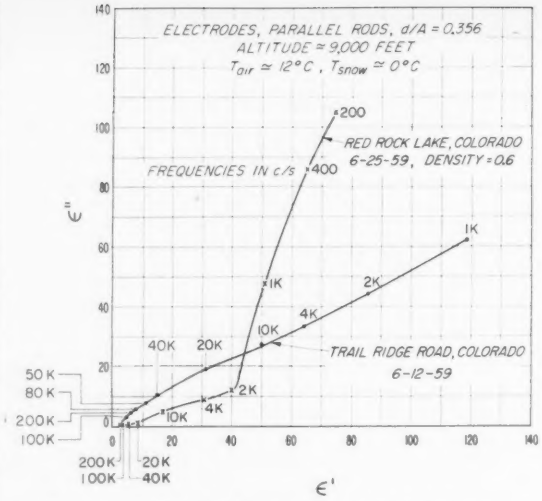


FIGURE 8. *Electrical properties of spring drift snow.*
Compact and wet.

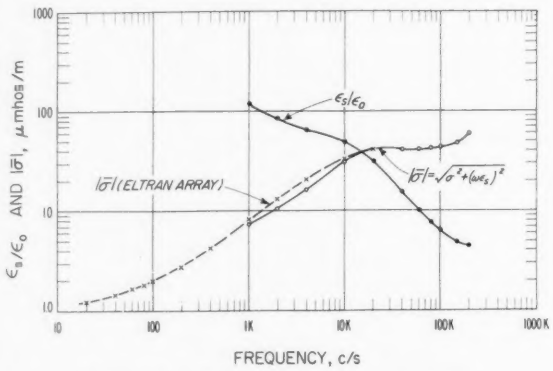


FIGURE 9. *Electrical properties of wet snow.*
Trail Ridge Road, Colorado, June 12, 1959.

departures from pure ice are in general agreement with observations made by Kuroiwa [8] and Devaney [9]. It should be pointed out that the rather large values of ϵ'' which result from an increase in conductivity are probably due to the presence of impurities in the snow and possible interface (overvoltage) effects at the particles [4]. The effect of impurities was verified by Kuroiwa by means of chemical analysis of the snow and by an experiment in which he produced hoarfrost crystals from chemically pure water and found that the Argand diagram was a semicircle as observed with pure ice although the dielectric constant was, of course, much lower because of the decreased density.

4.2. Athabasca Glacier, Columbia Icefields, Canada

Observations made on the Athabasca glacier near Jasper in Alberta, Canada, produced the Argand diagram shown in figure 10. It is seen from this

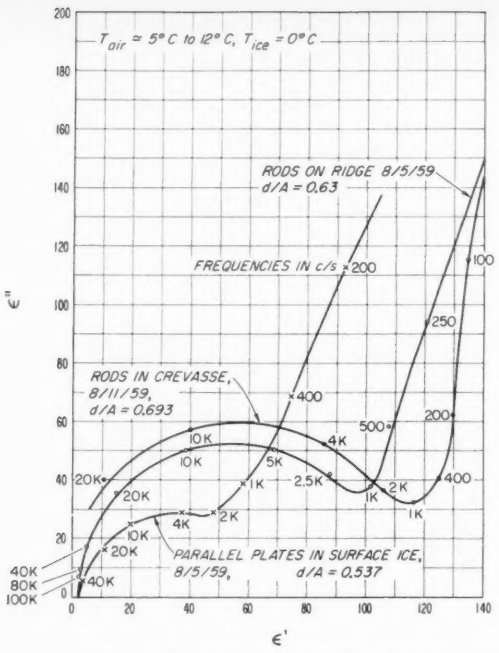


FIGURE 10. *Electrical properties of glacial ice.*
Athabasca glacier, Alberta, Canada.

figure that the general circular form expected for pure ice having a simple dielectric relaxation is obtained for frequencies down to approximately 2 kc. Below this the ϵ'' values increase quite rapidly indicating an appreciable amount of conductivity.

It should be noted that the radius of these Argand diagrams is greater than that predicted for 0° C pure ice from figure 4. This general type of departure was also observed by Auty and Cole [6] who explained the discrepancy on the basis of electrode interface effects. They indicate that this effect disappears at larger electrode spacings which, for the equipment they employed, would be a d/A ratio of 29. The smallest ratio of d/A which they employed was in the order of 3.3. In our experiments d/A varied from 0.36 to 0.69. A comparison of the magnitude of the complex conductivity as obtained by bridge and four-electrode Eltran measurements, figure 11, shows close agreement. This may be due to the fact that the rods were surrounded by a very thin interface of water. These sets of observations were obtained in the same general area on the glacier and would indicate that the large observed values of ϵ' and ϵ'' are probably real. It should be pointed out that some dispersion was found between various observations in different areas on the glacier due possibly to cracks and crevasses.

The parallel plate observations of figure 10 were made near the surface of the ice which was rather porous and wet at the time of measurement. The rods on the ridge were placed in the same general area as the parallel plate measurements, but being 4 ft long are expected to give results typical for ice several

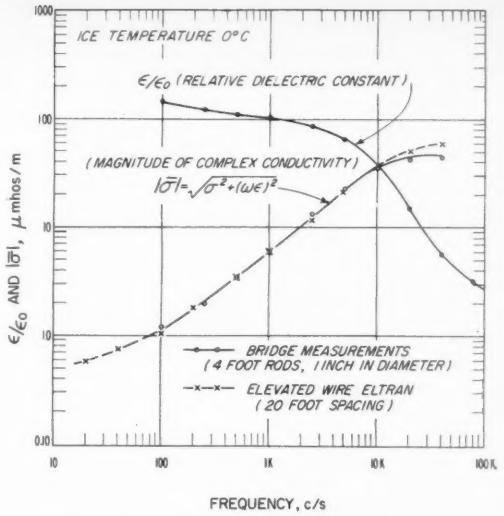


FIGURE 11. *Electrical properties of glacial ice.*
Athabasca glacier, Alberta, Canada, Aug. 5, 1959.

feet below the surface which was observed to be much more solid than the surface ice. The third curve for rods in a crevasse represents measurements made some 30 ft below the surface of the glacier where the ice was less disturbed by wind action, and in addition, it was expected to be slightly colder than the surface ice. This reduction in temperature with depth is indicated by the fact that rods in or near the surface caused melting around them while rods a hundred feet or so below the surface would freeze in place.

Figure 12 shows $|\bar{\sigma}|$ as a function of frequency for three different electrode spacings on the Athabasca

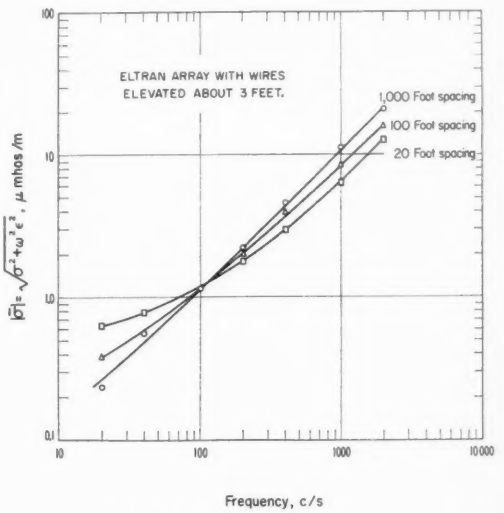


FIGURE 12. *Magnitude of complex conductivity vs. frequency at three electrode spacings.*
Athabasca glacier, Alberta, Canada, Aug. 1959.

glacier. It is interesting to note the reversal in magnitude at the 100-cps region. Similar data is shown in figure 13 where $|\bar{\sigma}|$ is plotted as a function of electrode spacing at different frequencies. It is obvious from these results that the properties of the glacier change with depth and that this change has an inverse effect at frequencies above and below the 100-cps region.

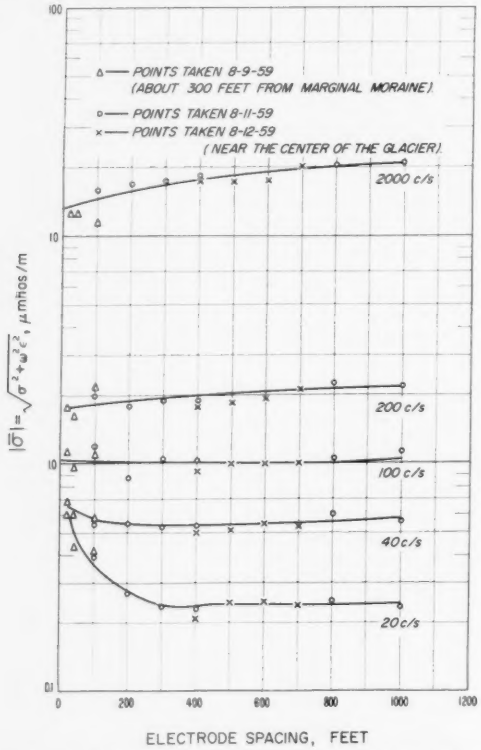


FIGURE 13. *Variations of magnitude of complex conductivity with electrode spacing.*
Athabasca glacier, Alberta, Canada.

The magnitude of the complex conductivity, $|\bar{\sigma}|$, obtained by means of the bridge measurements is shown in figure 14, and it is interesting to observe that on the surface $|\bar{\sigma}|$ is greater below 200 cps and less above this frequency than is found for the subsurface measurements. This general behavior is in agreement with the observations of figures 12 and 13 when the variation of effective penetration with stake spacing is considered. This effective physical depth for the various electrode spacings is discussed in considerable detail by Wait [3].

It is instructive to observe the manner in which the conductivity, admittivity, and loss tangent vary as a function of frequency. A typical example taken from the Athabasca glacier for 0° C glacial ice is shown in figure 15 where it can be seen that the loss tangent increases below 1 kc rather than continuing to decrease as it does for pure ice.

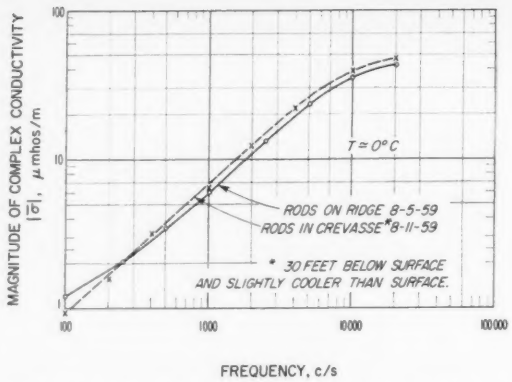


FIGURE 14. *Magnitude of complex conductivity for surface and subsurface ice.*
Athabasca glacier, Alberta, Canada.

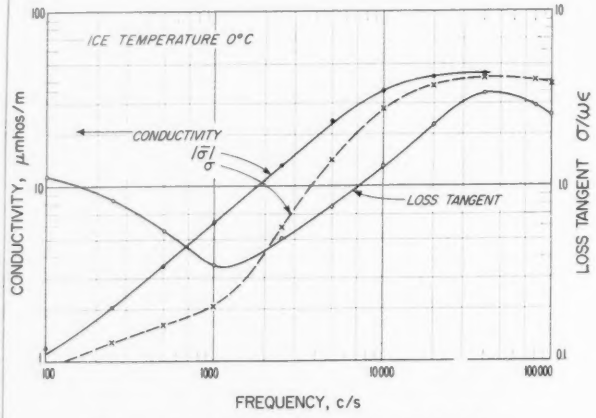


FIGURE 15. *Electrical properties of glacial ice.*
Athabasca glacier, Alberta, Canada, Aug. 5, 1959.

5. Conclusions

It is apparent from our limited observations that the electrical properties of snow and glacial ice are appreciably different than those of pure ice. The greatest difference appears in the region below 2 kc. The conductivity, σ, of glacial ice appears to be approaching a constant in the vicinity of 0.5 to 3 μ mhos/m for frequencies of 200 cps and lower. For pure ice, σ appears to be continuing to decrease with decreasing frequency and at 100 cps, $\sigma \cong 0.01$ μ mhos/m at $-0.1°$ C. Typical values of σ at 20 kc are 30 μ mhos/m for pure ice, and 36 μ mhos/m for the ice of the Athabasca glacier.

The close agreement between the values of complex conductivity magnitude obtained with the bridge measurements and four electrode methods would appear to indicate that electrode surface effects have not greatly influenced the results reported here. For future bridge measurements it

does, however, appear desirable to employ several rod spacings including d/A ratios up to 10 or more. For general propagation studies and wave-tilt calculations, the values of σ and ϵ given for ice should in most cases be used in the stratified earth formulas of Wait [10] rather than in the simple relation for an infinitely deep layer. Typical values of σ for the material below the ice may vary from 0.1 to 10 millimhos/m. Surface measurements near the Athabasca glacier yielded values essentially independent of frequency of $\sigma \sim 2$ millimhos/m for soil in the creek beds, and ~ 0.3 millimhos/m for a rocky ridge.

The authors acknowledge the assistance obtained from: J. R. Wait, for many valuable suggestions during the course of this work; E. C. Bolton and R. W. Plush, for assistance in preparation of equipment and obtaining data; Henri Vetter and D. R. Watt, for assisting in the measurements on the Athabasca Glacier: G. D. Garland, whose interest made the measurements on the glacier possible; F. J. Weyl, for his interest and encouragement during the course of this work; and Mrs. Winifred Mau, for her assistance during the preparation of this manuscript.

6. References

[1] J. R. Wait and A. M. Conda, Pattern of an antenna on a curved lossy surface, IRE Trans. PGAP **AP–6,** 348 (1958).
[2] A. D. Watt, E. L. Maxwell, and E. H. Whelan, Low-frequency propagation paths in Arctic areas, J. Research NBS **63D,** 99 (1959).
[3] J. R. Wait and A. M. Conda, On the measurement of ground conductivity at VLF, IRE Trans. PGAP **AP–6,** 273 (1958).
[4] J. R. Wait (Ed.), Overvoltage research and geophysical applications (Pergamon Press, New York, N.Y., 1959).
[5] N. E. Dorsey, Properties of ordinary water-substance (Reinhold Publishing Corp., New York, N.Y., 1940).
[6] R. P. Auty and R. H. Cole, Dielectric properties of ice and solid D_2O, J. Chem. Phys. **20,** 1309 (1952).
[7] J. G. Powles, A calculation of the static dielectric constant of ice, J. Chem. Phys. **20,** 1302 (1952).
[8] D. Kuroiwa, The dielectric property of snow (Inst. Low Temp. Sci., Hokkaido Univ., Soygcoro, Japan) Proc. IGU, Rome (1954).
[9] T. E. Devaney (private communication).
[10] J. R. Wait, Propagation of radio waves over stratified ground, Geophysics **18,** 416 (1953).

(Paper 64D4–69)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Half-Wave Cylindrical Antenna in a Dissipative Medium: Current and Impedance[1]

Ronold King [2] and Charles W. Harrison [3]

(January 7, 1960)

An integral equation for the distribution of current along a cylindrical antenna in a conducting dielectric is derived. It is shown that the boundary conditions for an antenna in such a medium are formally the same as for an antenna in free space. The equation is solved for the current I and the driving-point impedance Z by means of a technique that achieves sufficiently high accuracy in the leading terms of an iteration procedure so that the higher-order terms do not need to be evaluated. Moreover, these leading terms consist only of trigonometric functions with complex coefficients. The electromagnetic field in the infinite dissipative medium may be computed relatively easily since the current in the antenna is expressed in such simple terms.

A numerical analysis is made to determine the properties of an antenna with an electrical length of one-half wavelength in the medium with conductivity σ and relative dielectric constant ϵ_r. Universal curves are given of $I/\sqrt{\epsilon_r}$ with $\sigma/\omega\epsilon_0\epsilon_r$ as the parameter and of $Z\sqrt{\epsilon_r}$ with $\sigma/\omega\epsilon_0\epsilon_r$ as the variable in the range $0 \leq \sigma/\omega\epsilon_0\epsilon_r \leq 0.4$. A table of numerical values of the impedance is given for media such as an isotropic ionosphere, dry salt, dry earth, wet earth, and lake water.

1. Introduction

A study of the properties of antennas in dissipative media is an interesting theoretical problem that may have practical significance. For example, it appears to be possible to make certain diagnostic measurements in the area of fusion reactors that depend on a knowledge of the impedance of an antenna in an ionized medium. Specifically, it may be possible to determine ionization densities in high-temperature plasmas in this manner. In one phase of a contemplated field operation known as "Plowshare," it is expected that certain information concerning the operation of a nuclear device buried in a salt mine can be obtained with a low-frequency radar and a dipole antenna embedded in salt. Such measurements require a knowledge of the characteristics of antennas immersed in such a medium as well as of signal attenuation and dispersion.

A knowledge of the shift of impedance of a stub-type missile antenna as it enters an ionized cloud (such as the ionosphere, the aurora australis or borealis, or one of the Van Allen belts) makes it possible to determine transmission loss due to antenna mismatch.

Generally speaking, if the distribution of current along an antenna in a dissipative medium is known the input admittance is also available, and it is a straightforward process, at least in principle, to calculate the electromagnetic field at any point in the medium. From the field the signal attenuation in the medium may be determined. This is of importance in telemetry work, and also in the theory of communications between submerged submarines.

A theory is developed in this paper for a perfectly conducting, symmetrical, center-driven antenna of finite radius immersed in an infinite medium of moderate attenuation. Attention is directed specifically to half-wave antennas. For example, the present method of analysis permits the calculation of the impedance of this antenna when buried in "poor earth" for which the relative dielectric constant is $\epsilon_r=7$, and the conductivity is $\sigma=10^{-3}$ mho/m. The theory also applies to a base-driven quarter-wave radiator oriented at right angles to an infinite perfectly conducting plane. It is of practical importance that the impedance of a stub-type

[1] Contribution from Sandia Corp., Albuquerque, N. Mex.

[2] Consultant to the Sandia Corp. and Gordon McKay Professor of Applied Physics, Harvard University.

[3] Member of the Scientific Staff, Sandia Laboratory.

544328—60——5

antenna protruding from a missile differs negligibly from the impedance of the same antenna when mounted on an infinite ground plane, provided that the dimensions of the carrying vehicle are not too small in terms of the wavelength. Note also that the impedance of an antenna in a homogeneous dissipative medium of finite extent is essentially the same as in an infinite medium provided the boundary is sufficiently far removed from the antenna that the electromagnetic field at the interface is reduced to a small value owing to absorption.

2. Basic Electromagnetic Theory

Maxwell's equations in a homogeneous dissipative medium are

$$\text{curl } \mathbf{E}=-j\omega\mathbf{B} \tag{1}$$

$$\text{curl } \mathbf{B}=j\omega\mu\xi\mathbf{E} \tag{2}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic vectors, respectively, and $\omega=2\pi f$ is the radian frequency. The constants $\mu=1/\nu$; σ; and ϵ are the absolute permeability, conductivity, and dielectric constant of the medium; and ξ is the complex dielectric factor given by

$$\xi=\epsilon\left(1-j\frac{\sigma}{\omega\epsilon}\right). \tag{3}$$

The suppressed time dependence is exp $(j\omega t)$.

The vector potential $\mathbf{A}$ and scalar potential ϕ are defined by

$$\text{curl } \mathbf{A}=\mathbf{B} \tag{4}$$

$$\text{div } \mathbf{A}=-j\omega\mu\xi\phi \tag{5}$$

$$-\text{grad } \phi=\mathbf{E}+j\omega\mathbf{A} \tag{6}$$

where (5) is the Lorentz condition. The elimination of ϕ from (5) and (6) leads to the following relation for the electric field in terms of the vector potential:

$$\mathbf{E}=-\frac{j}{\omega\mu\xi}(\text{grad div } \mathbf{A}+k^2\mathbf{A}) \tag{7}$$

where

$$k^2=\omega^2\mu\xi=(\beta-j\alpha)^2. \tag{8}$$

The numerical evaluation of k may be expedited by noting that if $p=\sigma/\omega\epsilon$,

$$k=\beta-j\alpha=\omega\sqrt{\mu\epsilon}(\sqrt{1-jp})=\omega\sqrt{\mu\epsilon}\{f(p)-jg(p)\} \tag{9}$$

where $f(p)=\cosh\ (\frac{1}{2}\ \sinh^{-1}p)$ and $g(p)=\sinh(\frac{1}{2}\ \sinh^{-1}p)$. Tables of the functions $f(p)$ and $g(p)$ are available in the literature [1].[4]

When $p^2<<4$, $k=\beta-j\alpha\approx\omega\sqrt{\mu\epsilon}-j\frac{\sigma}{2}\sqrt{\mu/\epsilon}$. Note that in this case $p=2\alpha/\beta=\frac{\sigma}{\omega\epsilon}$, $\alpha^2/\beta^2<<1$.

The boundary conditions between the homogeneous dissipative medium, region 1, and the perfectly conducting cylindrical antenna, region 2, are

$$\left.\begin{aligned}\epsilon_1\hat{n}_1\cdot\mathbf{E}_1&=-\eta_{1f}-\eta_{2f}\\ \hat{n}_1\times\mathbf{E}_1&=0\end{aligned}\right\} \tag{10}$$

$$\left.\begin{aligned}\nu_1\hat{n}_1\times\mathbf{B}_1&=\mathbf{K}_{2f}\\ \hat{n}_1\times\mathbf{B}_1&=0\end{aligned}\right\} \tag{11}$$

[4] Figures in brackets indicate the literature references at the end of this paper.

since $\mathbf{E}_2=\mathbf{B}_2=0$ in the perfect conductor. $\hat{n}_1$ is a unit vector directed normally out of region 1, $\eta_{1f}+\eta_{2f}$ is the total surface density of charge on the boundary, and $\mathbf{K}_{2f}$ is the current density on the surface of the antenna.

A surface equation of continuity [2] that applies to the thin layer containing the surface charges and currents may be expressed as follows:

$$\text{div } \mathbf{K}_2+j\omega(\eta_{1f}+\eta_{2f})-\hat{n}_1\cdot\sigma\mathbf{E}_1=0. \tag{12}$$

It may be satisfied in two parts. On the perfectly conducting surface,

$$\text{div } \mathbf{K}_{2f}+j\omega\eta_{2f}=0 \tag{13}$$

and on the adjacent surface of the dissipative medium,

$$j\omega\eta_{1f}-\hat{n}_1\cdot\sigma\mathbf{E}_1=0. \tag{14}$$

Thus η_{2f} is the surface density of charge associated with the surface density of current $\mathbf{K}_{2f}$ along the perfect conductor, and η_{1f} is the surface density of charge related to the component of the volume density of current

$$J_{1f}=\sigma E_1 \tag{15}$$

that leaves the conductor normally to enter the medium in which it is immersed.

The substitution of the value of η_{1f} from (14) into (10) and (11) brings the boundary conditions into the following symmetrical form:

$$\left.\begin{aligned}\xi\hat{n}_1\cdot\mathbf{E}_1&=-\eta_{2f}\\ \hat{n}_1\times\mathbf{E}_1&=0\end{aligned}\right\} \tag{16}$$

$$\left.\begin{aligned}\nu\hat{n}_1\times\mathbf{B}_1&=-\mathbf{K}_{2f}\\ \hat{n}_1\cdot\mathbf{B}_1&=0.\end{aligned}\right\} \tag{17}$$

3. Vector Potential on the Surface of the Antenna

Let the axis of the antenna (fig. 1) coincide with the z-axis of a cylindrical system of coordinates r, θ, z. Let its ends be at $z=\pm h$; its cylindrical surface at $r=a$. It is assumed that

$$\left.\begin{aligned}\beta a&\ll 1\\ a&\ll h\end{aligned}\right\} \tag{18}$$

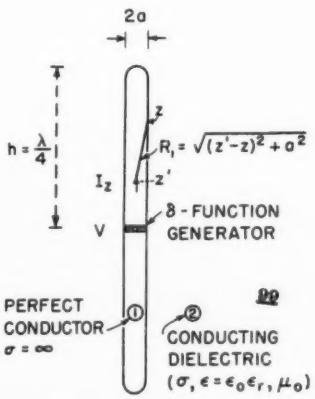


FIGURE 1. *Half-wave dipole in a conducting medium.*

(where β is defined by (9)). These conditions and an appropriate method of excitation that generates only a z-component of current in the antenna make it possible to set

$$\mathbf{K}_{2f}=\hat{z}K_{2z}. \tag{19}$$

The substitution of (4) in (17) with $\hat{n}_1=-\hat{r}$, gives

$$\nu\hat{n}_1\times\mathbf{B}_1=\nu\hat{n}_1\times\text{curl }\mathbf{A}=\nu\hat{r}\times\text{curl }\mathbf{A}=-\hat{z}K_{2z}. \tag{20}$$

If $A_r=A_\theta=\dfrac{\partial A}{\partial\theta}=0$, (20) becomes

$$\nu\hat{r}\times\hat{\theta}\left(-\frac{\partial A_z}{\partial r}\right)=\hat{z}K_{2z} \tag{21}$$

or

$$\left(\frac{\partial A_z}{\partial r}\right)_{r=a}=\left(\frac{\partial A_z}{\partial n}\right)_{r=a}=\frac{K_{2z}}{\nu}. \tag{22}$$

It is well known that the equation satisfied by the z-component of the vector potential in an infinite homogeneous dissipative medium is

$$\nabla^2A_z+k^2A_z=0. \tag{23}$$

Its solution may be obtained by applying Green's theorem in the symmetric form,

$$\int_\tau[u\nabla^2v-v\nabla^2u]d\tau'=\int_z\left[u\frac{\partial v}{\partial n}-v\frac{\partial u}{\partial n}\right]d\sigma' \tag{24}$$

with $u=A_z$ and $v=G$, where G is the free space Green's function that satisfies the auxiliary equation

$$\nabla^2G+k^2G=-\delta(R_d). \tag{25}$$

In (25), $R_d=|\mathbf{R}-\mathbf{R}'|$ is the distance between the point located by the end of the vector $\mathbf{R}'$ where G is calculated and a point source at the end of the vector $\mathbf{R}$. The Dirac delta function satisfies the relations

$$\delta(R_d)=0,\ R_d\neq0 \tag{26}$$

$$\int_\tau\delta(R_d)d\tau\begin{cases}1 \text{ if } \tau \text{ includes } R_d=0\\ 0 \text{ if } \tau \text{ excludes } R_d=0.\end{cases} \tag{27}$$

The solution of (25) that vanishes at infinity is

$$G=\frac{e^{-jkR_d}}{4\pi R_d}. \tag{28}$$

Equation 24 now becomes

$$\int_\tau\left[G(\nabla^2A_z+k^2A_z)-A_z(\nabla^2G+k^2G)\right]d\tau'=\int_{\Sigma_1+\Sigma_2}\left(G\frac{\partial A_z}{\partial n}-A_z\frac{\partial G}{\partial n}\right)d\sigma' \tag{29}$$

where Σ_2 is the surface of a great sphere, and Σ_1 is any parallel pair of surfaces within Σ_2 that enclose surfaces across which the functions G and A_z or their normal derivatives are discontinuous. In the case at hand the only such surface is the envelope of the perfectly conducting antenna of radius a, and the only one of the four functions that is discontinuous across it is $\partial A_z/\partial n$ as given in (22). It follows that with (23), (25), and (27) together with (22), (29) becomes

$$\int_\tau A_z\delta(\mathbf{R}-\mathbf{R}')d\tau'=\frac{1}{\nu}\int_{\Sigma_1}K_{2z}(z')Gd\sigma'+\int_{\Sigma_2}\left(G\frac{\partial A_z}{\partial n}-A_z\frac{\partial G}{\partial n}\right)d\sigma'. \tag{30}$$

The volume integral reduces to $A_z=A_z(\mathbf{R})$. Since k is complex with negative imaginary part, both A_z and G vanish at infinity. Accordingly, the integral over Σ_2 approaches zero as Σ_2 recedes to infinity. With $I_z(z')=2\pi aK_{2z}(z')$, (30) becomes

$$A_z(\mathbf{R})=\frac{1}{4\pi\nu}\int_h I_z(z')dz'\int_{-\pi}^{\pi}\frac{e^{-jkR_s}}{R_s}\frac{d\theta'}{2\pi} \tag{31}$$

where

$$R_s=\sqrt{(z-z')^2+\left(2a\sin\frac{\theta'}{2}\right)^2} \tag{32}$$

If $h\gg a$, as is assumed, an adequate approximation of (31) is

$$A_z(\mathbf{B})=\frac{1}{4\pi\nu}\int_{-h}^{h}I_z(z')\frac{e^{-jkR_1}}{R_1}dz' \tag{33}$$

where

$$R_1=\sqrt{(z-z')^2+a^2}. \tag{34}$$

The vector potential A_z is defined in terms of the surface current $I_z=2\pi aK_{2z}$ on the perfect conductor. This current vanishes at the ends of the antenna if the end surfaces are neglected or if it is a tube and I_z is the sum of the currents on inner and outer surfaces. That is

$$I_z(\pm h)=0. \tag{35}$$

From (13), the surface current satisfies the equation of continuity

$$\frac{dK_{2z}}{dz}+j\omega\eta_{2f}=0, \tag{36}$$

or when multiplied by $2\pi a$,

$$\frac{dI_z}{dz}+j\omega q_2=0 \tag{37}$$

where $q_2=2\pi a\eta_{2f}$ is the surface charge per unit length associated with the current along the conductor. The axial current is associated with q_2, and the radial diffusion current in the medium with η_{1f} in the equation of continuity

$$j\omega\eta_{1f}-(\sigma E_{1r})_{r=a}=0. \tag{38}$$

Thus, the physical picture is a sheet of current $I_z=2\pi aK_{2z}$ on the surface of the perfect conductor that is directed only axially and decreases by charging the surface. A portion of this charge remains for a part of the period in the axial standing wave, and a portion leaves radially into the adjacent imperfectly conducting or ionized medium. This is contained in the equation

$$\frac{dI_z}{dz}+j\omega q_2+j\omega q_1-2\pi a\sigma E_{1r}=0 \tag{39}$$

where $q_1=2\pi a\eta_{1f}$.

4. Integral Equation for the Current

The integral equation for the current in a perfectly conducting antenna of length $2h$ and radius a, center-driven at $z=0$ by an idealized delta-function generator with electromotive force V may be derived easily from (23) and (33) with $\nabla^2=\partial^2/\partial z^2$. It is [3]

$$\int_{-h}^{h}I_z(z')\frac{e^{-jkR_1}}{R_1}dz'=-j\frac{4\pi}{\zeta}\left\{C_1\cos kz+\frac{V}{2}\sin k|z|\right\}. \tag{40}$$

In this relation $R_1=\sqrt{(z-z')^2+a^2}$, C_1 is a constant evaluated from the boundary condition (35), and

$$\zeta=\sqrt{\frac{\mu}{\xi}}=\sqrt{\frac{j\omega\mu}{\sigma+j\omega\epsilon}}. \tag{41}$$

Equation (40) is valid in an infinite homogeneous dissipative medium. In order that the present analysis may parallel a recent development in the theory of linear antenna arrays [4], an alternative form of (40) is desired. It is

$$\int_{-h}^{h} I_z(z')\left\{\frac{e^{-jkR_1}}{R_1}-\frac{e^{-jkR_h}}{R_h}\right\}dz'=-j\frac{4\pi}{\zeta}\left\{C_1\cos kz+\frac{V}{2}\sin k|z|+U\right\} \tag{42}$$

where

$$U=-j\frac{\omega\mu}{4\pi k}\int_{-h}^{h} I_z(z')\frac{e^{-jkR_h}}{R_h}dz'=-j\frac{\omega}{k}A_z(h) \tag{43}$$

and

$$R_h=\sqrt{(h-z')^2+a^2}.$$

When $z=h$, (42) becomes

$$C_1\cos kh+\frac{V}{2}\sin kh+U=0 \tag{44}$$

so that

$$C_1=-\left\{\frac{\frac{1}{2}V\sin kh+U}{\cos kh}\right\}. \tag{45}$$

This value of C_1 may be substituted into (42) to obtain

$$\int_{-h}^{h} I_z(z')K_D(z,z')dz'=j\frac{4\pi k}{\omega\mu\cos kh}\{\tfrac{1}{2}V\sin k(h-|z|)+U(\cos kz-\cos kh)\} \tag{46}$$

where

$$K_D(z,z')=\frac{e^{-jkR_1}}{R_1}-\frac{e^{-jkR_h}}{R_h}=e^{-\alpha R_1}\frac{e^{-j\beta R_1}}{R_1}-e^{-\alpha R_h}\frac{e^{-j\beta R_h}}{R_h}. \tag{47}$$

Equation (46) is the integral equation for the current along a symmetrical center-driven antenna immersed in an infinite dissipative medium of arbitrary attenuation. The only restrictions that have been imposed are given by (18).

Since a delta-function generator is employed to drive the antenna, the driving point current, which defines the input admittance, is rigorously infinite because there is included an infinite admittance across the input terminals given by [5]

$$\underset{\infty}{Y}=\underset{\infty}{G}+j\omega\underset{\infty}{C}=\frac{4a}{\zeta}\ln\left\{\frac{1}{\left|z'+\frac{h'}{2}\right|}\right\} \tag{48}$$

as $z'\to-h'/2$ where $z'=z-h$, $h'=2h$. The conductance and the capacitance are infinite since the knife edges between the idealized halves of the antenna are separated zero distance and are in contact with the dissipative medium. However, if (40) or (46) is solved in terms of continuous functions by the usual method of iteration, the driving point current actually obtained is essentially that maintained by the delta-function generator *minus this infinite gap current.* This is the desired approximate solution for the current in a practical antenna driven by an actual transmission line if suitable end corrections are provided.

5. Approximate Solution of the Integral Equation for the Current

Let a solution of (46) be sought specifically for the half-wave dipole defined by $\beta h=\pi/2$. The procedure followed is described in the literature [4]. In brief, it involves the separation of the total current along the antenna and the vector potential on its surface into several parts and the proper association of these to permit the introduction of several expansion functions that are sensibly constant along the antenna. Each of these functions is defined as the ratio of a particular component of the vector potential to its associated current. Several integral equations are obtained in this way which may be iterated independently of one another. It has been found that the sum of *only the leading* trigonometric terms of these component currents with suitably defined complex coefficients yields the total current and the driving point impedance with satisfactory accuracy.

The solution of (46) is begun by studying the variation with respect to z of the functions [6]

$$\Phi_v(z)=\int_h^h \sin k(h-|z'|)K_D(z, z')dz', \tag{49}$$

and

$$\Phi_u(z)=\int_h^h (\cos kz'-\cos kh)K_D(z, z')dz'. \tag{50}$$

Now

$$\sin k(h-|z|)=\cos \beta z \cosh\left\{\alpha h\left(1-\frac{|z|}{h}\right)\right\}-j \sin \beta|z| \sinh\left\{\alpha h\left(1-\frac{|z|}{h}\right)\right\}, \tag{51}$$

and

$$\cos kz-\cos kh=\cos \beta z \cosh\left\{\alpha h\cdot\frac{z}{h}\right\}-i\left[\sinh \alpha h-\sin \beta z\ \beta z \sinh\left\{\alpha h\cdot\frac{z}{h}\right\}\right] \tag{52}$$

when $\beta h=\pi/2$. These quantities do not differ significantly in general form from those with $\alpha=0$, and the additional terms (which vanish when $\alpha=0$) are not sufficiently great to dominate provided the condition $\alpha h\leq 1$ is imposed. Actually, when $\alpha=\beta$, $\alpha h=1.57$. The slightly more severe condition,

$$(\alpha h)^2<<2 \text{ or } \alpha h\leq 0.3 \tag{53}$$

permits the following approximations:

$$\cosh \alpha h\approx 1, \sinh\approx \alpha h, e^{-\alpha R}\approx 1-\alpha R. \tag{54}$$

It follows that

$$\sin k(h-|z|)\approx \beta z-j\alpha h\left(1-\frac{|z|}{h}\right)\sin \beta|z| \tag{55}$$

and

$$\cos kz-\cos kh\approx \cos \beta z-j\alpha h\left(1-\frac{z}{h}\sin \beta z\right). \tag{56}$$

If the approximations

$$\left(1-\frac{|z|}{h}\right)\sin \beta|z|\approx \cos \beta z+\sin \beta|z|-1 \tag{57}$$

and

$$1-\frac{z}{h}\sin \beta z\approx \cos \beta z \tag{58}$$

are made, it follows that

$$\sin k(h-|z|)\approx(1-j\alpha h)\cos \beta z+j\alpha h(1-\sin \beta|z|) \tag{59}$$

and

$$\cos kz-\cos kh\approx(1-j\alpha h)\cos \beta z. \tag{60}$$

The approximations (57) and (58) are used throughout the analysis of the half-wave dipole antenna in a dissipative medium. They may be shown to be very good when $\beta h = \pi/2$ by direct graphical comparison. Since $e^{-\alpha R_1} \approx 1 - \alpha R_1$, $K_D(z,z')$, as defined by (47), may be written

$$K_D(z, z') = K_d(z, z') - \alpha K_e(z, z') \tag{61}$$

where

$$K_d(z, z') = \frac{e^{-j\beta R_1}}{R_1} - \frac{e^{-j\beta R_h}}{R_h} \tag{62}$$

and

$$K_e(z, z') = e^{-j\beta R_1} - e^{j\beta R_h}. \tag{63}$$

With (59) to (61), inclusive, (49) becomes

$$\begin{aligned}\Phi_v(z) \approx (1 - j\alpha h) \int_h^h \cos \beta z \{K_d(z, z') - \alpha K_e(z, z')\} dz' \\ + j\alpha h \int_h^h (1 - \sin \beta |z|) \{K_d(z, z') - \alpha K_e(z, z')\} dz' \\ \approx \left(1 - j\frac{\alpha\lambda}{4}\right)\left\{C_a\left(\frac{\lambda}{4}, z\right) - C_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right) - \frac{\alpha}{\beta}\left[\frac{\pi}{2}(1+j) - j\right]\cos \beta z\right\} \\ + j\alpha\frac{\lambda}{4}\left\{E_a\left(\frac{\lambda}{4}, z\right) - E_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right) - S_a\left(\frac{\lambda}{4}, z\right) + S_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right)\right\}.\end{aligned} \tag{64}$$

Terms of order $\alpha^2 h/\beta$ and $\alpha^2 h^2$ have been neglected. The particular values of the general functions C_a, E_a and S_a[5] required in this application are given later. In a similar manner (50) becomes

$$\begin{aligned}\Phi_u(z) \approx (1 - j\alpha h)\left\{\int_{-h}^h \cos \beta z' K_d(z, z') dz' - \alpha \int_{-h}^h \cos \beta z' K_e(z, z') dz'\right\} \\ \approx \left(1 - j\alpha\frac{\lambda}{4}\right)\left\{C_a\left(\frac{\lambda}{4}, z\right) - C_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right) - \frac{\alpha}{\beta}\left[\frac{\pi}{2}(1+j)\right]\cos \beta z\right\}.\end{aligned} \tag{65}$$

It is known [4] that $C_a(\lambda/4, z) - C_a\lambda/4, \lambda/4)$ varies approximately as $\cos \beta z$ in both its real and imaginary parts. It follows that $\Phi_u(z)$ does also. Accordingly

$$\Phi_u(z) \approx \Phi_u(0) \cos \beta z \tag{66}$$

where

$$\Phi_u = \Phi_u(0) = \left(1 - j\alpha\frac{\lambda}{4}\right)\left\{C_a\left(\frac{\lambda}{4}, 0\right) - C_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right) - \frac{\alpha}{\beta}\left[\frac{\pi}{2}(1+j) - j\right]\right\}. \tag{67}$$

A comparison of (64) and (65) shows that

$$\Phi_v(z) \approx \Phi_u(z) + \Phi_w(z) \tag{68}$$

where

$$\Phi_w(z) = j\alpha\frac{\lambda}{4}\left\{E_a\left(\frac{\lambda}{4}, z\right) - E_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right) - S_a\left(\frac{\lambda}{4}, z\right) + S_a\left(\frac{\lambda}{4}, \frac{\lambda}{4}\right)\right\}. \tag{69}$$

---

[5] These functions $C_a(h, z)$ and $S_a(h, z)$ and $E_a(h, z)$ are defined as follows (see ref. [3], Sec. 19): $C_a(h, z) = \int_0^h L \cos \beta z' dz'$, $S_a(h, z) = \int_0^h L \sin \beta z' dz'$, $E_a(h, z) = \int_0^h L dz'$, where $L = \{(e^{-j\beta R_1}/R_1) + (e^{-\beta R_2}/R_2)\}$, $R_1 = \sqrt{(z-z')^2 + a^2}$, $R_2 = \sqrt{(z+z')^2 + a^2}$. They may be expressed in terms of tabulated functions.

Since it is readily verified that $\Phi_w(z)$ varies like $(1-\sin\beta|z|)$ in both its real and imaginary parts, it follows that

$$\Phi_v(z)\approx\Phi_u(0)\cos\beta z+\Phi_w(0)(1-\sin\beta|z|). \tag{70}$$

In (70),

$$\Phi_w=\Phi_w(0)=j\alpha\frac{\lambda}{4}\left\{E_a\left(\frac{\lambda}{4},0\right)-E_a\left(\frac{\lambda}{4},\frac{\lambda}{4}\right)-S_a\left(\frac{\lambda}{4},0\right)+S_a\left(\frac{\lambda}{4},\frac{\lambda}{4}\right)\right\}. \tag{71}$$

From (68) or (70),

$$\Phi_v\approx\Phi_u+\Phi_w. \tag{72}$$

This completes the investigation of the functions $\Phi_v(z)$ and $\Phi_u(z)$.

With (59) to (61), and the approximation $\cos kh\approx j\alpha h$, (46) may be expressed as follows:

$$\int_{-h}^{h}I_{v1}(z')\{K_d(z,z')-\alpha K_e(z,z')\}dz'+\int_{-h}^{h}I_{v2}(z')\{K_d(z,z')-\alpha K_e(z,z')\}dz'+\int_{-h}^{h}I_u(z')\{K_d(z,z')$$

$$-\alpha K_e(z,z')\}dz'=\frac{4\pi(\beta-j\alpha)}{\omega\mu\alpha h}\left\{\left[\frac{1}{2}V+U\right][1-j\alpha h]\cos\beta z+j\frac{1}{2}V\alpha h(1-\sin\beta|z|)\right\} \tag{73}$$

where $I_z(z')$ has been replaced by

$$I_z(z')=I_{v1}(z')+I_{v2}(z')+I_u(z'). \tag{74}$$

The integrals in (73) may be written as follows:

$$\left.\begin{aligned}\int_{-h}^{h}I_{v1}(z')\{K_d(z,z')-\alpha K_e(z,z')\}dz'&=\psi_{v1}I_{v1}(z)-D_{v1}(z)\\ \int_{-h}^{h}I_{v2}(z')\{K_d(z,z')-\alpha K_e(z,z')\}dz'&=\psi_{v2}I_{v2}(z)-D_{v2}(z)\\ \int_{-h}^{h}I_u(z')\{K_d(z,z')-\alpha K_e(z,z')\}dz'&=\psi_uI_u(z)-D_u(z)\end{aligned}\right\} \tag{75}$$

where

$$\psi_i=\int_{-h}^{h}g_i(z,z')\{K_d(z,z')-\alpha K_e(z,z')\}dz' \tag{76}$$

and

$$D_i=\int_{-h}^{h}[I_i(z')-I_i(z)g_i(z,z')][K_d(z,z')-\alpha K_e(z,z')]dz' \tag{77}$$

where the subscript i stands for v_1, v_2, or u.

An inspection of (73) and the $\Phi(z)$ functions previously studied, shows that a proper association of the components of the vector potential and the current is achieved if the following choice of distribution functions is made:

$$\left.\begin{aligned}g_{v1}(z,z')=g_u(z,z')&=\frac{\cos\beta z'}{\cos\beta z}\\ \text{and}\qquad g_{v2}(z,z')&=\frac{1-\sin\beta|z'|}{1-\sin\beta|z|}\end{aligned}\right\}. \tag{78}$$

Then, with (76),

$$\left.\begin{aligned}\psi_{v1}=\psi_u&=\frac{\Phi_u}{1-j\alpha h}\\ \text{and}\qquad \psi_{v2}&=\frac{\Phi_w}{j\alpha h}\end{aligned}\right\}. \tag{79}$$

This pairing of vector potentials and currents insures that to a high degree of approximation. $\psi_{v1}\sim A_{v1}(z)/I_{v1}(z)$, $\psi_{v2}\sim A_{v2}(z)/I_{v2}(z)$, and $\psi_u\sim A_u(z)/I_u(z)$. $A_{v1}(z)$, $A_{v2}(z)$, and $A_u(z)$ are components of the vector potential calculated from the currents $I_{v1}(z)$, $I_{v2}(z)$, and $I_u(z)$, respectively. It follows that all difference integrals, defined by (77), are small.

The substitution of (75) into (73) gives

$$I_{v1}(z)\psi_{v1}+I_{v2}(z)\psi_{v2}+I_u(z)\psi_u=\frac{4\pi(\beta-j\alpha)}{\omega\mu\alpha h}\left\{\left(\frac{1}{2}V+U\right)(1-j\alpha h)\cos\beta z+j\frac{1}{2}V\alpha h(1-\sin\beta|z|)\right\}+D_{v1}(z)+D_{v2}(z)+D_u(z). \quad (80)$$

Since $\psi_{v1}=\psi_u$ and the components of the vector potential contributed by $I_{v1}(z)$ and $I_u(z)$ vary as $\cos\beta z$, (80) may be separated into two parts and these iterated separately. Thus

$$I_{v1}(z)+I_u(z)=\frac{4\pi(\beta-j\alpha)}{\omega\mu\alpha h\psi_u}\left\{\frac{1}{2}V+U\right)(1-j\alpha h)\cos\beta z\Big\}+\frac{1}{\psi_u}\left\{D_{v1}(z)+D_u(z)\right\}. \quad (81)$$

$$I_{v2}(z)=j\frac{4\pi(\beta-j\alpha)}{\omega\mu\alpha h\psi_{v2}}\left\{\frac{1}{2}V\alpha h(1-\sin\beta|z|)+\frac{D_{v2}(z)}{\psi_{v2}}\right\}. \quad (82)$$

The leading terms may be substituted into the difference integrals to obtain first order corrections to the currents, if desired.

The total current is

$$I_z(z)=\frac{2\pi(\beta-j\alpha)V}{\omega\mu\alpha h\psi_u}\left\{\left(1+\frac{2U}{V}\right)(1-j\alpha h)\cos\beta z+j\alpha h(1-\sin\beta|z|)\frac{\psi_u}{\psi_{v2}}+\cdots\right\}. \quad (83)$$

The ratio $2U/V$ may be determined from (43) with the substitution of $I_z(z)$ as given by (83), i.e.,

$$U=-j\frac{\omega\mu}{4\pi(\beta-j\alpha)}\int_{-h}^{h}I_z(z')\left\{\frac{e^{-j\beta R_{1h}}}{R_{1h}}-\alpha e^{-j\beta R_{1h}}\right\}dz'$$

$$=-j\frac{V}{2\alpha h\psi_u}\left\{\left(1+\frac{2U}{V}\right)(1-j\alpha h)\psi_c(h)+j\alpha h\frac{\psi_s(h)\psi_u}{\psi_{v2}}\right\} \quad (84)$$

where

$$\psi_c(h)=\int_{-h}^{h}\cos\beta z'\left\{\frac{e^{-j\beta R_{1h}}}{R_{1h}}-\alpha e^{-j\beta R_{1h}}\right\}dz'\approx C_a\left(\frac{\lambda}{4},\frac{\lambda}{4}\right)+j\alpha h \quad (85)$$

and

$$\psi_s(h)=\int_{-h}^{h}(1-\sin\beta|z'|)\left\{\frac{e^{-j\beta R_{1h}}}{R_{1h}}-\alpha e^{-j\beta R_{1h}}\right\}dz'\approx E_a\left(\frac{\lambda}{4},\frac{\lambda}{4}\right)-S_a\left(\frac{\lambda}{4},\frac{\lambda}{4}\right)+j\frac{\alpha}{\beta} \quad (86)$$

The solution of (84) for $2U/V$ yields

$$\frac{2U}{V}=\frac{\frac{\psi_s(h)\psi_u}{\psi_{v2}}-\left(1+\frac{j}{\alpha h}\right)\psi_c(h)}{\psi_u+\psi_c(h)\left(1+\frac{j}{\alpha h}\right)}. \quad (87)$$

The substitution of (87) into (83) gives the following final expression for the leading terms of the current:

$$I_z(z)=-j\frac{2\pi(\beta-j\alpha)V}{\omega\mu\psi_{v2}}\left\{T\left(\frac{\lambda}{4}\right)\cos\beta z+\sin\beta|z|-1\right\}. \quad (88)$$

The corresponding driving point impedance is

$$Z(k)=\frac{V}{I_z(0)}=j\,\frac{\omega\mu\psi_{v2}}{2\pi(\beta-j\alpha)\left\{T\left(\frac{\lambda}{4}\right)-1\right\}}. \tag{89}$$

In (88) and (89),

$$T\left(\frac{\lambda}{4}\right)=\frac{\left(1-j\,\frac{\pi}{2}\,\frac{\alpha}{\beta}\right)\left(\psi_{v2}+\psi_s\left(\frac{\lambda}{4}\right)\right)}{\left(1-j\,\frac{\pi}{2}\,\frac{\alpha}{\beta}\right)\psi_c\left(\frac{\lambda}{4}\right)-j\frac{\pi}{2}\,\frac{\alpha}{\beta}\,\psi_u}. \tag{90}$$

The use of approximte forms of the C_a, S_a, and E_a functions [6], leads to the following formulas. The final numerical results in (91) apply to a half-wave dipole with $h/a=75$ or $a/\lambda=1/300$.

$$\psi_u=2\,\sinh^{-1}\frac{h}{a}-2.357-j\,0.633-\frac{\alpha}{\beta}\,(1.571+j\,0.571) \tag{91a}$$

$$=7.663-j\,0.633-\frac{\alpha}{\beta}\,(1.571+j\,0.571 \qquad \text{for } \frac{h}{a}=75 \tag{91b}$$

$$\psi_{v2}=4\,\sinh^{-1}\frac{h}{a}-2\,\sinh^{-1}\frac{2h}{a}-1.747-j\,0.384 \tag{92a}$$

$$=6.885-j\,0.384 \qquad \text{for } \frac{h}{a}=75 \tag{92b}$$

$$\psi_c=0.709-j\,1.219+j\,\frac{\pi\alpha}{2\beta} \tag{93}$$

$$\psi_s=2\,\sinh^{-1}\frac{2h}{a}-2\,\sinh^{-1}\frac{h}{a}-1.219-j\,0.709+j\,\frac{\alpha}{\beta} \tag{94a}$$

$$=0.167-j\,0.709+j\,\frac{\alpha}{\beta} \qquad \text{for } \frac{h}{a}=75. \tag{94b}$$

6. Distribution of Current

The current as given in (88) involves the parameter $\beta/\omega\mu=(1/\zeta_0)\sqrt{\epsilon_r/\mu_r}\,f(p)$ and the dimensionless ratio $p=2\alpha/\beta=\sigma/\omega\epsilon$ that characterize the properties of the medium in which the antenna is immersed. If attention is restricted to nonmagnetic materials and values of $p\leq 0.5$, it follows that $f(p)\approx 1$[6] so that $\beta/\omega\mu=\sqrt{\epsilon_r}/\zeta_0$. In this case the two relevant quantities are $\sqrt{\epsilon_r}$ and $\sigma/\omega\epsilon_0\epsilon_r$. It follows that a *normalized* current may be defined that is a function of $2\alpha/\beta$ or of $\sigma/\omega\epsilon_0\epsilon_r$ *alone* insofar as the properties of the medium are concerned. That is,

$$\frac{I_z}{V\sqrt{\epsilon_r}}=\frac{-j2\pi\left(1-j\frac{\alpha}{\beta}\right)}{\zeta_0\psi_{v2}}\left\{T\left(\frac{\lambda}{4}\right)\cos\beta z+\sin\beta|z|-1\right\} \tag{95}$$

where $\zeta_0=\sqrt{\mu_0/\epsilon_0}\approx 120\pi$ ohms and it is assumed that $\sigma/\omega\epsilon_0\epsilon_r\leq 0.5$. With the current expressed in the form $I_z=I_z'+jI_z''=I_z\epsilon^{j\theta_I}$ the family of curves shown in figure 2 has been evaluated from (95) for the distribution of the real and imaginary parts of the normalized current along a half-wave dipole. The parameter is $\sigma/\omega\epsilon_0\epsilon_r$. The corresponding curves of the amplitude and phase are shown in figure 3. Note that these curves apply to an antenna for which $h/a=75$ and $\beta h=\pi/2$ with $\beta=\omega\sqrt{\mu\epsilon}$ the phase constant appropriate to the dielectric medium. The actual

[6] When $p=0.5$, $f(p)=1.029$ which differs from unity by less than 3 percent.

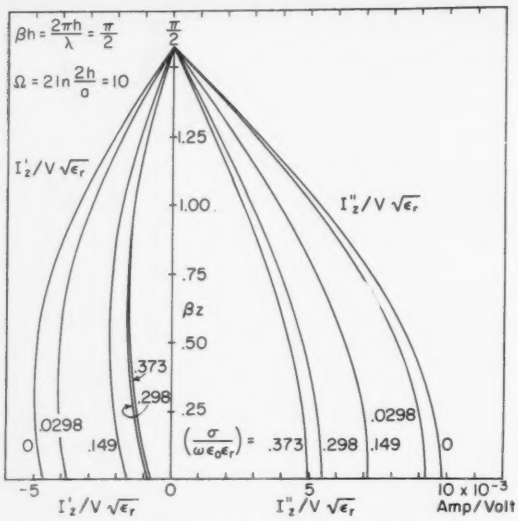


FIGURE 2. *Current* $I_z = I_z'' + jI_z'$ *in normalized form for a half-wave dipole in a medium with conductivity* σ *and dielectric constant* $\epsilon = \epsilon_0\epsilon_r$; I_z'' *is in phase with the driving voltage* V.

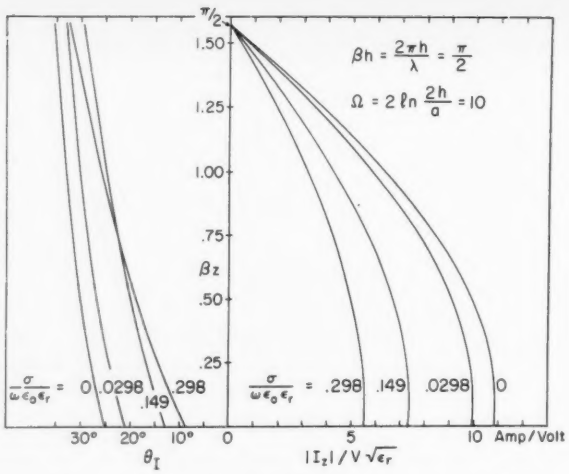


FIGURE 3. *Current in figure 2 in the form* $I_z = I_z e^{j\theta_I}$.

currents per input volt are obtained from these figures by multiplying the normalized values with $\sqrt{\epsilon_r}$. The curves indicate that when an antenna is immersed in a medium of given dielectric constant, the amplitude of the current diminishes with increasing conductivity of the medium if the same driving voltage is applied. The inductive lag of the input current also decreases as σ is made greater, but there is a rise in the relative phase change outward along the antenna. That is, the current near the end of the antenna lags behind the current at the driving point by an angle that increases with σ.

7. Admittance and Impedance

The admittance, $Y = G + jB$, and the impedance, $Z = R + jX$, are most conveniently expressed in the normalized forms, $Y/\sqrt{\epsilon_r}$ and $Z\sqrt{\epsilon_r}$. The former is obtained directly from (95) with $z = 0$; evidently, $Z\sqrt{\epsilon_r} = 1/(Y\sqrt{\epsilon_r})$. The normalized quantities, $G/\sqrt{\epsilon_r}$, $-B/\sqrt{\epsilon_r}$ and $R\sqrt{\epsilon_r}$, $X\sqrt{\epsilon_r}$ are shown graphically in figure 4. In the range $(\sigma/\omega\epsilon_0\epsilon_r) \leq 0.5$ a single curve is obtained for each of these quantities for all values of σ and ϵ_r. It is seen that when an antenna is immersed in a medium with a given dielectric constant for which $2h = \lambda/2$ in the medium, its input resistance increases almost linearly with increasing conductivity of the medium while its reactance first decreases to a shallow minimum, then increases slightly. The actual admittance of a half-wave antenna when immersed in a medium with a given dielectric constant is obtained from figure 4 by multiplying the normalized conductance and susceptance by $\sqrt{\epsilon_r}$. Similarly, the impedance is obtained from the normalized resistance and reactance when divided by $\sqrt{\epsilon_r}$. Numerical values of actual resistances and reactances for a range of dielectric constants and conductivities are listed in the table 1.

It is important to bear in mind that the data contained in figure 4 and table 1 always apply to an antenna that is exactly a half wavelength long *in the medium* with the specified dielectric constant and conductivity. That is, the electrical half-length is $\beta h = \pi/2$. Note, however, that if an antenna has the half-length $h = \lambda/4$ referred to the wavelength in the medium, its half-length referred to free space is $h = \lambda_0/4\sqrt{\epsilon_r}$. In other words, the electrical length $\beta_0 h$ in free space corresponding to an electrical length $\beta h = \pi/2$ in the medium is $\beta_0 h = \beta h/\sqrt{\epsilon_r}$ when $\mu = \mu_0$. If an antenna that is adjusted in length so that $\beta h = \pi/2$ in a medium with constants

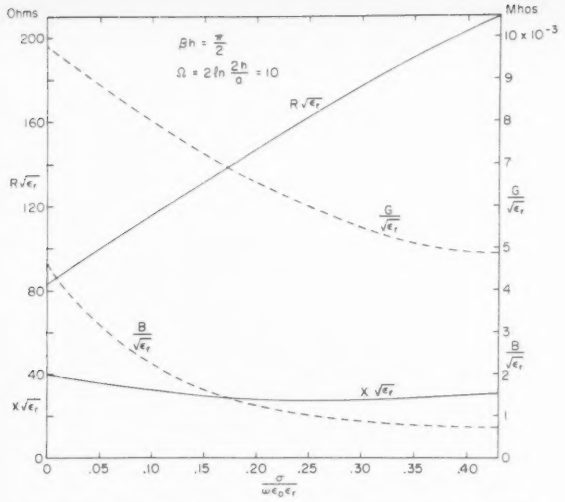


FIGURE 4. *Impedance and admittance of a half-wave dipole when immersed in a medium with conductivity* σ *and dielectric constant* $\epsilon = \epsilon_0\epsilon_r$; $\beta = \omega\sqrt{\mu\epsilon} = 2\pi/\lambda$.

TABLE 1. *Impedances of antennas for which* $\beta h = \pi/2$ *when immersed in a dissipative medium;* h/a=75, f=6 *Mc*

$Z(k)$ in ohms	Medium	σ in mho/m	$\sigma/\omega\epsilon_0\epsilon_r$	ah
$\epsilon_r = 0.1$				
283.3+j126.6		0	0	
283.3+j126.3		4.9×10^{-8}	0.00146	
286.7+j125.6		1.5×10^{-7}	.00448	
294.6+j118.3		10^{-6}	.0298	
310.5+j114.2		1.5×10^{-6}	.0448	
$\epsilon_r = 0.665$				
102.05+j49.01		0	0	
102.08+j48.99	Ionosphere	3.26×10^{-7}	0.00146	
103.4 +j48.7		10^{-6}	.00448	
120.4 +j44.3		10^{-5}	.0448	
235.6 +j35.1		8×10^{-5}	.358	0.274
$\epsilon_r = 1.0$				
83.2+j40.0	Free space	0	0	
93.1+j37.4		10^{-5}	0.0298	
132.6+j29.8		5×10^{-5}	.149	
178.5+j28.0		10^{-4}	.298	.230
$\epsilon_r = 6.6$				
32.4+j15.6		0	0	
32.8+j15.4		10^{-5}	0.00452	
40.3+j13.4	Dry salt	1.34×10^{-4}	.066	
65.8+j10.7		6×10^{-4}	.272	.205
$\epsilon_r = 7.0$				
78.5+j11.7	Dry earth	10^{-3}	.426	.322
$\epsilon_r = 10.0$				
26.3+j12.6		0	0	
29.4+j11.8		10^{-4}	0.0298	
41.9+j9.44		5×10^{-4}	.149	
55.9+j8.72	Wet earth	10^{-3}	.298	.230
$\epsilon_r = 80.0$				
9.30+j4.47		0	0	
9.40+j4.44	Distilled water	10^{-4}	0.00373	
10.7 +j4.11	Lake water	10^{-3}	.0373	
21.9 +j3.24	Lake water	10^{-2}	.373	.285

ϵ_r and σ is removed to free space, its impedance changes both because the parameters that characterize the medium are different and because its electrical length is changed. This is illustrated in figure 5 which gives R and X for an antenna that has been removed to free space from a medium with relative dielectric constant ϵ_r in which its electrical length was $\pi/2$.

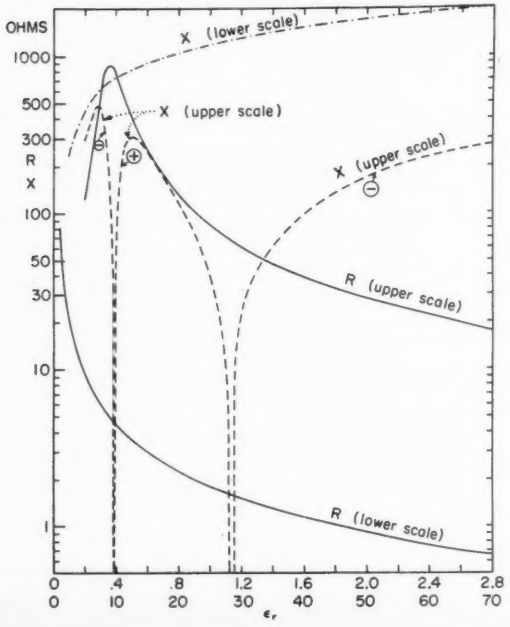


FIGURE 5. R *and* X *of dipole in free space with half-length* h *(for which* $\beta h = \omega h\sqrt{\mu_0\epsilon_0\epsilon_r}$ *when the antenna is immersed in a dielectric with relative dielectric constant* ϵ_r *so that* $h = \left(\frac{\lambda_0}{4\sqrt{\epsilon_r}}\right)$.

8. Special Cases

A number of special cases are of interest. The following data apply to an antenna for which $\beta h=\pi/2$ in the specified medium and $h/a=75$.

Case I. Antenna in free space. $\epsilon_r=1$, $\sigma=0$, so that $\alpha=0$ and $\beta=\beta_0=2\pi/\lambda_0$.

$$Z(\beta_0)=83.2+j\,40.0 \text{ ohms.} \tag{96a}$$

The King-Middleton second order theory gives [7]

$$Z(\beta_0)=86.5+j\,41.7 \text{ ohms.} \tag{96b}$$

It is noteworthy that the new approach to the analysis of cylindrical antennas [4] yields results that agree so closely with those of a complicated, twice-iterated solution (that is known to be in good agreement with experiment when proper account is taken of transmission-line end and coupling effects [8]), since only the leading trigonometric terms in the current distribution are used in obtaining (96a). This is a consequence of properly choosing the trigonometric functions and carefully adjusting their relative complex values.

Case II. Split missile sounding rocket. According to Nicolet [9] at an altitude of 100 m (where the temperature is 210° K) the maximum collision frequency ν [7]$=1.1\times10^5$ sec^{-1}. At this same altitude, DiTaranto and Lamb [10] anticipate an electron density $N=1.5\times10^{11}$ electrons m^3. The standard relations for the dielectric constant and conductivity of the ionosphere, neglecting the earth's magnetic field are

$$\epsilon=\epsilon_0\left\{1-\frac{Ne^2}{\epsilon_0 m(\nu^2+\omega^2)}\right\} \tag{97a}$$

and

$$\sigma=\frac{Ne^2\nu}{m(\nu^2+\omega^2)}, \tag{97b}$$

where $\epsilon_0=8.85\times10^{-12}$ f/m, e is the charge of the electron ($e=-1.60\times10^{-19}$ coulomb), and m is the mass of the electron ($m=9.11\times10^{-31}$ kg). With the above values for N and ν, a frequency of 6 *Mc*, $\epsilon=0.6649\epsilon_0$ and $\sigma=3.26\times10^{-7}$ mho/m, the attenuation constant of the medium is $\alpha=7.534\times10^{-5}$ nepers/m and the phase constant is $\beta=0.1025$ radians/m. The impedance of the center-driven half-wave dipole in this medium is

$$Z(k)=102.1+j49.0 \text{ ohms.} \tag{98}$$

In air the electrical half-length of this antenna is $\beta_0 h=1.926$, its actual half length $h=15.3$m, and its impedance is $Z(\beta_0)=200+j203$ ohms. On comparing $Z(k)$ with $Z(\beta_0)$, it is clear that the electron density of the *normal ionosphere* can easily be determined with a half-wave dipole antenna used as a probe at 6 Mc.

Case III. Antenna in a Plasma. If the electron density is 4.05×10^{11} electrons/m^3 and the collision frequency is $\nu=6.15\times10^3$ sec^{-1} it follows that $\epsilon=0.1\epsilon_0$, $\sigma=4.9\times10^{-8}$ mho/m, $\beta=0.040$ radians/m, $\alpha=2.92\times10^{-3}$ nepers/m, and the impedance of the half-wave dipole is

$$Z(k)=263.3+j126.3 \text{ ohms.} \tag{99}$$

The half-length of the antenna at $f=6$ Mc is $h=\pi/2\beta=39.3$ m.

Case IV. Antenna Buried in Salt. The measured properties of the salt at 6 Mc are $\epsilon_r=6.6$ and $\sigma=1.34\times10^{-4}$ mho/m. It follows that $\beta=0.323$ radians/m, $\alpha=9.85\times10^{-3}$ nepers/m, and the impedance of the half-wave dipole is

$$Z(k)=40.3+j13.5 \text{ ohms.} \tag{100}$$

[7] ν should not be confused with $\nu=1/\mu$. ν is retained as the symbol for collision frequency in order to agree with standard notation.

Since

$$\beta h=\frac{\pi}{2},\ h=\frac{\pi}{2\beta}=\frac{\pi}{2\times 0.323}=4.86 \text{ m}.$$

Case V. Antenna in "poor earth" for which $\epsilon_r=7$ and $\sigma=10^{-3}$ mho/m. At 6 Mc, $\beta=0.34$ radians/m, and $\alpha=0.0697$ nepers/m, and

$$Z(k)=78.47+j11.66 \text{ ohms}. \tag{101}$$

It should be mentioned that in this case $\alpha h=\alpha\pi/2\beta=0.322$. This value of αh is near the limit of validity of the approximations introduced in the analysis and contained in the condition $\alpha h\leq 0.3$.

Case VI. Antenna in a highly conducting medium. If the antenna is immersed in a highly conducting medium, the formula (88) for the distribution of current does not apply owing to the restriction $\alpha h\leq 0.3$. However, irrespective of the precise form of the distribution of current along the antenna, the formula (89) for the impedance should approach the correct limit, $Z(k)\to 0$, as $\sigma\to\infty$. When $\sigma/\omega\epsilon$ is large compared with unity

$$\beta-j\alpha=(1-j)\sqrt{\frac{\omega\sigma\mu}{2}} \tag{102}$$

as $\sigma\to\infty$, $\beta=\alpha\to\infty$, but $\alpha/\beta=1$. An examination of the ψ_u, ψ_{v2}, ψ_c, and ψ_s functions shows that they remain finite. Since $\beta-j\alpha$ appears in the denominator of (89), and all other functions are well behaved, $Z(k)\to 0$ as $\sigma\to\infty$.

9. Conclusion

Although the present analysis has been applied specifically to an antenna that is a half wavelength long when immersed in a conducting dielectric, and the conductivity of the dielectric has been limited to moderate values by the condition $\alpha h\leq 0.3$, the method is applicable to other lengths and higher conductivities. A study of antennas for which $\alpha h\leq 0.3$ but $\beta h=\pi$ is in progress. The electrically short antenna defined by $\beta h<1$ is also under investigation.

Since the properties of a two-wire line immersed in a conducting dielectric are understood, the extension of the present study to the folded dipole is readily accomplished with the method of symmetrical components. Details are reserved for a subsequent paper.

10. References

[1] R. W. P. King, Electromagnetic engineering, app. 2, pp. 510 to 518 (McGraw-Hill Book Co., Inc., New York, N.Y., 1945).

[2] Ronold King, Electromagnetic engineering, Ch. I, Sec. 24, pp. 77 to 81, especially equation 13 (McGraw-Hill Book Co., New York, N.Y., 1956).

[3] R. W. P. King, Theory of linear antennas, Ch. II (Harvard University Press, Cambridge, Mass., 1956).

[4] Ronold King, Linear arrays: Currents, impedances, and fields, I. (Harvard University, Cambridge, Mass., May 1, 1959) Trans. IRE **AP-7,** 5440 (1959).

[5] Tai Tsun Wu and Ronold King, Driving point and input admittance of linear antennas, J. Appl. Phys. **30,** 74 (1959).

[6] The logic of this step is made clear by a study of sections 2 and 4 of the paper, The cylindrical antenna: current and impedance, by Ronold King and David Middleton, Quart. Appl. Math. 4 (July 1946). An alternative reference is sec. 20 in ch. II of R. W. P. King, Theory of linear antennas (Harvard University Press, Cambridge, Mass., 1956).

[7] R. W. P. King, Theory of linear antennas, p. 168 (Harvard University Press, Cambridge, Mass., 1956).

[8] R. W. P. King, Theory of linear antennas, ch. II, sec. 38 (Harvard University Press, Cambridge, Mass., 1956).

[9] Marcel Nicolet, Collision frequency of electrons in the terrestrial atmosphere, Phys. Fluids **2,** 95 (1959).

[10] R. A. DiTaranto and J. J. Lamb, The space environment—a preliminary study, Elec. Manuf. (Gage Publishing Co., Oct. 1958).

10.1. Additional References

A. Baños, Jr., and J. P. Wesley, The horizontal electric dipole in a conducting half-space, pt. I (1953), pt. II (1954), Univ. Calif. Marine Phys. Lab., Scripps Inst. Oceanog. (unpublished).

C. W. Harrison, Jr., and D. H. Denton, Jr., Impedance of a dipole in an ionized medium—I, Sandia Corp., New Mexico (1959) (unpublished).

C. W. Harrison, Jr., and D. H. Denton, Jr., Impedance of a dipole in an ionized medium—II, Sandia Corp., New Mexico (1959) (unpublished).

C. W. Harrison, Jr., Calculation of the impedance of a cylindrical antenna in an ionized medium of low attenuation factor by means of tables of the exponential integral for complex arguments, Sandia Corp., New Mexico (1959) (unpublished).

M. S. Macrakis, Cylindrical antenna immersed in a dissipative medium (input impedance), Cruft Lab., Harvard Univ. (1957) (unpublished).

R. K. Moore, The theory of radio communication between submerged submarines, Ph. D. Thesis, Graduate School of Cornell Univ. (June 1951) (unpublished).

C. T. Tai, Radiation of a Hertzian dipole immersed in a dissipative medium, Cruft Lab., Harvard Univ. (1947) (unpublished).

C. T. Tai, On radiation and radiating systems in the presence of a dissipative medium, Cruft Lab., Harvard Univ. (1949) (unpublished).

J. R. Wait and L. L. Campbell, The fields of an electric dipole in a semi-infinite conducting medium, J. Geophys. Research **58,** 21 (1953).

J. R. Wait, The radiation fields of a horizontal dipole in a semi-infinite dissipative medium, J. Appl. Phys. **24,** 958 (1953).

J. R. Wait and L. L. Campbell, Fields of dipoles in a semi-infinite conducting medium, J. Geophys. Research **58,** 21 to 28, 167 to 178 (1953).

J. R. Wait, Insulated loop antenna immersed in a conducting medium, J. Research NBS **59,** 133 (1957) RP2781.

J. R. Wait, Radiation from a small loop immersed in a semi-infinite conducting medium, Can. J. Phys. **37,** 672 (1959).

(Paper 64D4–70)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 4, July–August 1960

Preface to ELF Papers

The remaining papers in this issue (and the first three papers in the next) were presented at the Conference on the Propagation of Electromagnetic Waves at Extremely Low Frequencies (ELF). This conference was held at the Boulder Laboratories of the National Bureau of Standards on January 26, 1960, under the guidance of A. D. Watt, Assistant Chief of the Radio Communication and Systems Division.

Extremely low frequencies are defined here as the range from 3 kc down to about 1.0 cps, and they are well below those currently used in communications. Lightning discharges, however, produce considerable energy in this range, and the radiated fields have been used for studying the nature of lightning phenomena for many years at frequencies as low as 10 cps. Other natural sources of both a terrestrial and an extraterrestrial nature also radiate electromagnetic energy in this frequency range but usually to a lesser extent than lightning.

Since ELF signals have been observed to propagate with very low attenuation, it has been suggested from time to time that they would be useful for communications on a worldwide scale. Furthermore, because of certain magneto-ionic phenomena, these waves may penetrate the ionosphere and enable communication with space vehicles. This is particularly so in the vicinity of 3 kc where a "window" in the ionosphere appears to exist. Another aspect of ELF waves is that the wavelength is extremely large (being 3,000 km at 100 cps) and consequently, the signals may easily be diffracted around planetary objects. Such signals also penetrate with relatively small loss into conducting media such as rocks and soil. In fact, frequencies in this range have been used for many years in geophysical exploration.

JAMES R. WAIT, *Editor.*

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 64D, No. 4, July–August 1960

Some ELF Phenomena [1]

E. T. Pierce

(February 23, 1960)

Properties of the electric and magnetic fields in natural extremely low frequency (ELF) phenomena are briefly discussed. The ELF fluctuations in the electric field are then treated from two aspects; these are the electromagnetic changes associated with atmospherics and the electrostatic variations in atmospheric electricity. A final section attempts to integrate the general subject of ELF effects of natural origin.

1. Introduction

Any antenna arrangement for receiving signals of natural origin may be envisaged as responding either to the electric (E) field or to the magnetic (H) field. An example of the first instance would be a vertical whip, and of the second an electrostatically shielded loop antenna. The E field has three components $E=E_S+E_I+E_R$ and the H field two terms $H=H_I+H_R$.

The individual components with suffix R are normally described as "radiation" or "electromagnetic"; those with the I suffix are termed "induction," "magnetic," or "current"; while the "electrostatic" component has the suffix S. H_I and E_I are, of course, interrelated, as are E_R and H_R; also, in free space the S, I, and R terms are proportional respectively to $1/d^3$, $1/d^2$, and $1/d$, where d is the distance from the source to the receiver. Furthermore, if the source is considered to radiate at a discrete frequency, as in the case of an oscillating dipole, then

$$\left.\begin{aligned} E_S:E_I:E_R &= 1:\frac{\omega d}{c}:\frac{\omega^2 d^2}{c^2} \\ H_I:H_R &= 1:\frac{\omega d}{c} \end{aligned}\right\}$$

where ω is the angular frequency concerned and c is the velocity of light. It is evident from these relations that for extremely low frequencies less than say 1 kc, the electrostatic component is dominant in the electric field even at considerable distances from the source. However, it should be remembered that in actual practice propagation is not unrestricted as in free space but confined, typically between the earth and the lower ionosphere, and when boundary conditions have to be applied the relations among the different field components are inevitably modified.[2]

Many aspects of ELF phenomena have been studied in the past. Two of these are considered in the succeeding sections, and some final discussion then attempts to integrate the whole field of ELF effects.

2. The "Slow Tail"

The radiation fields due to distant lightning or "atmospherics" as they are more commonly termed, often contain both ELF (<1 kc) and VLF (ca 10 kc) components. This effect is basically due to the pronounced propagational attenuation at around 1 to 3 kc; this has been established both experimentally [1] and theoretically [2]. A typical waveform of distant origin is depicted in figure 1. It is of an atmospheric due to a discharge to earth. The 10-msec sweep is initiated by a precursor; then follows the disturbance, some 300 mv/m peak-to-peak, known variously as the "VLF section," the "oscillatory head," or the "main stroke"; and this is succeeded by the "ELF portion" or "slow tail." Watson-Watt, Herd, and Lutkin [3] were the first to make a detailed study of the ELF section; they noticed the important feature that the farther away the origin of the atmospheric the greater the time separation t between the oscillatory head and the slow tail. Hales [4] applied the waveguide theory of propagation between the earth and a sharply bounded lower ionosphere to the ELF results of Watson-Watt and his associates. Hales was primarily concerned with the quasi zero-order mode. Hepburn and Pierce [5] established experimentally the following empirical laws relating t, d, and $\tau/4$, the time occupied by the first quarter cycle of the slow tail:

$$\text{day}\begin{cases} \dfrac{\tau}{4}=500+0.23d \\ t=0.33d-350 \end{cases}$$

$$\text{night}\begin{cases} \dfrac{\tau}{4}=500+0.08d \\ t=0.13d-140. \end{cases}$$

[1] Contribution from AVCO Research and Advanced Development Div., Wilmington, Mass.; paper presented at Conference on the Propagation of ELF Radio Waves, Boulder, Colo., Jan. 25, 1960.

[2] See paper by J. R. Wait in this issue for such an analysis.

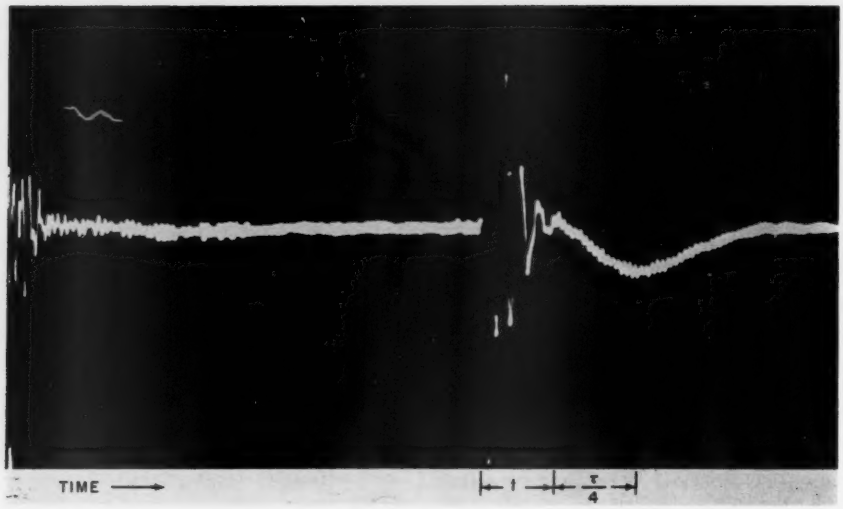


FIGURE 1. *A typical waveform.*

The units are kilometers for d and microseconds for t and τ.

Using the theory developed by Hales, Hepburn, and Pierce were able to deduce from their experimental results a range of pairs of solutions for h the height of the ionospheric stratum and σ its conductivity. Neither can be determined separately but the most likely solution can be assessed since there is other information available on the variation of electron density N and collisional frequency ν with height h, and therefore of σ, since this is mainly dependent on the ratio N/ν. The final result of Hepburn and Pierce was that the experimental observations by day indicated an ionospheric layer of height 65 km and conductivity 5×10^3 esu; at night the corresponding figures were 90 km and 10^4 esu. Hepburn [6] in a revised analysis later gave 70 km and 2×10^3 esu for day, and 90 km and 4×10^3 esu at night.

Liebermann [7] has also worked on slow tails. He derived his own theory which produced the relation

$$d=4h\ (\pi\sigma\tau)^{1/2}.$$

Here the units are kilometers for d, h and seconds for τ. Liebermann found that his experimental results fitted his theoretical relationship with values of h and σ of 60 and 10^4 esu by day; at night the figures were 90 and 5×10^4. There is obviously some discordance between the work of Liebermann and that of Hepburn and Pierce; it has not yet been resolved.

The question of the generation of slow tails is of some interest. It seems certain that many tails arise from a continuing current following the return stroke in a flash to earth. Pierce [8] in his observations of the electrostatic field-changes associated with close discharges has pointed out that some flashes seem to have a pronounced slow change of field after the return stroke; in other discharges this slow change appears to be absent. Thus there should be a tendency for some flashes to give large slow tails while others produce little or no ELF component. Hepburn, however, believes that the distribution is relatively continuous. In an effort to resolve this point a histogram (fig. 2) has been plotted of the ELF/VLF ratios for several hundred atmospherics recently recorded at Avco, Wilmington, Mass. The great majority of these, being winter atmospherics, tended to originate at comparable distances. The histogram indicates that the distribution is much more extended than would be expected from the smaller values of the ELF/VLF ratios, or, in other words, the contention that certain families of lightning flashes are extremely potent generators at ELF's seems to be supported. In this connection it is perhaps noteworthy that cloud discharges may produce strong ELF signals; if so, they would be expected to give an ELF to VLF ratio appreciably different from that in flashes to earth.

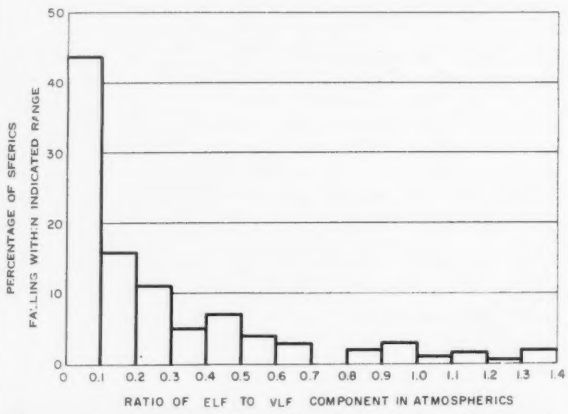


FIGURE 2. *Histogram illustrating the statistical behavior of the ELF/VLF ratio for atmospherics.*

3. ELF Signals in Atmospheric Electricity

In atmospheric electricity as compared with atmospherics, the ELF signals are predominantly electrostatic rather than electromagnetic. The work of Holzer and his school [9] represents a natural bridge in relating the two separate fields. It is well known that the electrostatic field changes accompanying a lightning flash often have durations approaching a second, and that violent fluctuations of the atmospheric potential gradient due to rain, point-discharge, space charge movements, and so on, occur in disturbed weather. These effects produce a spectrum of ELF phenomena extending both well above and well below 1 cps. Appreciable fluctuations at ELF's can, however, also occur in fair weather. Examples have been noted by Pierce [8] and one is reproduced in figure 3. The record is from a capillary electrometer; the total duration is 35 sec and the peak-to-peak amplitudes of the electrostatic fluctuations in potential gradient are of the order of 1 v/m.

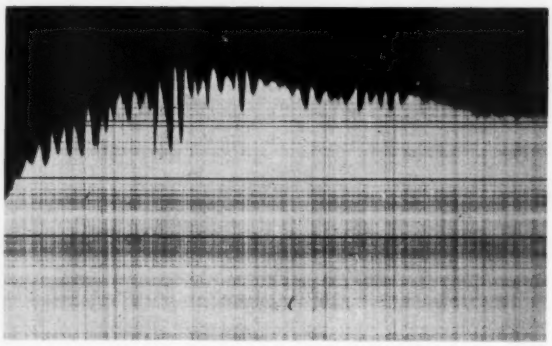


FIGURE 3. *ELF fluctuations in the fair weather electrostatic field.*

Large [10, 11] has made a detailed investigation of the behavior along the spectrum of variations in the electric field. He used experimental methods of recording simultaneously on two spaced aerials in order to separate the effects of local and of distant origin; the frequency range from 10 kc downwards was covered with broadband and with sharply tuned receivers. Large believes that because the natural relaxation or readjustment time of the lower atmosphere is of the order of 10^3 sec, such effects as changes in resistivity due to altering pollution will produce variations with frequencies less than 10^{-3} cps. From 10^{-3} to 1 cps Large finds that fair weather variations are due to the movement of space charges. These often travel with the wind at a height of 10 or 20 m. There is also a turbulent component which is well correlated with the diurnal variation of turbulence. Above 1 cps the situation is somewhat confused but the general picture seems to be the following. From 1 kc to 100 cps disturbances tend to be electromagnetic being the discrete slow tails well correlated with VLF atmospherics at 10 kc. Below 100 cps the correlation with 10 kc becomes less and less marked and the fluctuations not discrete but more and more continuous. This is for two reasons. Propagation at 100 cps is already better even than at 10 kc and is becoming increasingly good from 100 cps downwards; thus ever more distant sources of atmospherics are being detected, and the individual disturbances are therefore becoming increasingly frequent, that is, approaching a continuous noise. Also, with a decrease in frequency, the electrostatic fluctuations due to the movement of space charges are becoming more and more marked, and Large believes that they are the dominant factor for frequencies less than some 20 cps.

4. Discussion

Some kind of summary can now be made of the behavior of both the electric and magnetic fields for frequencies below 1 kc. As Large has shown the E field changes over from being dominantly electromagnetic and originating in distant lightning flashes, to being mainly electrostatic with a source in local movements of space charge. The changeover is at a frequency of about 20 cps. Below 20 cps the phenomena are almost entirely electrostatic, but at "extremely low frequencies indeed" the origin is no longer necessarily local. For instance with frequencies of the order of 10^{-4} to 10^{-5} cps the diurnal variations of local atmospheric pollution and of worldwide thunderstorm activity are both effective as regards changes in the atmospheric potential gradient. At 10^{-7} cps seasonal influences are entering and there are further possibilities of annual and longer duration variations.

The behavior of the H field is in some respects similar to that of the E field. Aarons [12] has pointed out that there is probably a changeover in dominance from the electromagnetic to the magnetic component at about 20 cps. Below this frequency the main features are caused by magnetic effects due to ionospheric and earth currents. Thus at the higher frequencies the E and H fields are both mostly electromagnetic and have a common origin in distant lightning flashes. Below about 20 cps the E field is predominantly electrostatic and the H field mostly magnetic; their sources, moreover, are entirely distinct. Therefore E and H measurements should correlate at the higher frequencies and be increasingly unrelated as frequency decreases. This is entirely the kind of behavior found when the work of Willis [13], who used loop antenna responding to the H field, is compared with the E measurements of Large.

5. References

[1] F. W. Chapman and R. C. V. Macario, Propagation of audio-frequency radio waves to great distances, Nature (London) **177,** 930 (1956).

[2] J. R. Wait, The attenuation vs frequency characteristics of VLF radio waves, Proc. IRE **45,** 768 (1957).

[3] R. A. Watson-Watt, J. F. Herd, and F. E. Lutkin, On the nature of atmospherics, Proc. Roy. Soc. [A] **162,** 267 (1937).

[4] A. L. Hales, A possible mode of propagation of the slow or tail components in atmospherics, Proc. Roy. Soc. [A] **193,** 60 (1948).
[5] F. Hepburn and E. T. Pierce, Atmospherics with very low-frequency components, Nature (London) **171,** 837 (1953).
[6] F. Hepburn, Atmospheric waveforms with very low frequency components below 1 kc/s known as slow tails, J. Atmos. Terr. Phys. **10,** 266 (1957).
[7] L. Liebermann, Extremely low-frequency electromagnetic waves. II. Propagation-properties, J. Appl. Phys. **27,** 1477 (1956).
[8] E. T. Pierce, Electrostatic field-changes due to lightning discharges, Quart. J. Roy. Met. Soc. **81,** 211 (1955).
[9] R. E. Holzer and O. E. Deal, Low audio-frequency electromagnetic signals of natural origin, Nature (London) **177,** 536 (1956).
[10] M. I. Large, Some investigations in atmospheric electricity, Dissertation (Cambridge Univ., Cambridge, England, 1957).
[11] M. I. Large and T. W. Wormell, Fluctuations in the vertical electric field, Recent Adv. Atmos. Elec., p. 603 (Pergamon Press, Ltd., New York, N.Y., 1958).
[12] J. Aarons, Low frequency electromagnetic radiation 10–900 cycles per second, J. Geophys. Research **61,** 647 (1956).
[13] H. F. Willis, Audio-frequency magnetic fluctuations, Nature (London) **161,** 887 (1948).

(Paper 64D4–71)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Mode Theory and the Propagation of ELF Radio Waves [1]

James R. Wait

(February 18, 1960)

The mode theory of propagation of electromagnetic waves at extremely low frequencies (ELF) (1.0 to 3,000 cycles per second) is treated in this paper. Starting with the representation of the field as a sum of modes, approximate formulas are presented for the attenuation and phase constants. Certain alternate representations of the individual modes are mentioned. These are used as a basis for describing the physical behavior of the field at large distances from the source, particularly near the antipode of the source. At the shorter distances, where the range is comparable to the wavelength, the spherical-earth mode series is best transformed to a series involving cylindrical wave functions. This latter form is used to evaluate the near field behavior of the various field components.

The effect of the earth's magnetic field is also evaluated using a quasi-longitudinal approximation. In general it is indicated that if the gyrofrequency is less than the effective value of the collision frequency, the presence of the earth's magnetic field may be neglected for ELF. When this condition is not met the attenuation may be increased somewhat. The influence of an inhomogeneous ionosphere is also briefly considered and, finally, the propagation of ELF pulses is treated. It is suggested that certain observed characteristics of ELF waveforms may be attributed to the inclination of the current channel in the lightning discharge.

1. Introduction

The propagation of electromagnetic waves at extremely low frequencies (ELF) has received considerable attention recently [1 to 5].[2] Unfortunately, the experimental data in this frequency range (1.0 cps to 3,000 cps) is still rather limited [8 to 15]. There is little doubt, however, that the bulk of the energy is transferred via a waveguide mode. The bounding surfaces of this waveguide are the ground and the lower edge of the ionospheric E region. In theoretical approaches to this problem, it is usually assumed that the ionosphere is homogeneous in its electrical properties. On comparing theoretical attenuation rates based on this assumption, it is apparent that the model must be refined to take account of the variation of the electron density with height. Such an extension to the mode theory has been made recently [16]. The influence of the earth's magnetic field is also expected to be important at ELF since the waves are reflected at a level where the collision frequency is of the same order as the gyrofrequency for electrons. The ionosphere is thus expected to be anisotropic. Unfortunately, the experimental data is not sufficient to confirm whether the anisotropy is significant.

Another feature of ELF propagation is that the distance between source and observer may be comparable to the wavelength. For example, at 300 cps the wavelength is 1,000 km. Many of the experiments are carried out at distances of this order and the computed attenuation rates must be considered with caution.

It is the purpose of this paper to survey the mode theory as it applies to ELF propagation and to develop several extensions. Many of the results are presented in graphical form as an aid to application.

---

[1] Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo. Paper presented at Conference on the Propagation of ELF Radio Waves, Jan. 25, 1960.

[2] Figures in brackets indicate the literature references at the end of this paper.

2. Basic Theoretical Model

The assumed theoretical model is taken to be a homogeneous conducting spherical earth of radius a surrounded by a concentric conducting ionospheric shell of inner radius $a+h$. It is convenient to introduce the usual spherical coordinate system (r, θ, ϕ). Using this model and assuming a vertical electric dipole source located at $\theta=0$, $r=a$, the expression for the vertical electric field, E_r, is given by an expression of the following form [4, 16]

$$E_r=\frac{Ids\eta}{2kha^2}\sum_{n=0}^{\infty}\delta_n\frac{\nu(\nu+1)}{\sin\nu\pi}P_\nu(-\cos\theta) \tag{1}$$

where

I—average current in the source dipole,
ds—length of source dipole,
η—intrinsic impedance of air space $\cong 120\pi$ ohms,
$P_\nu(-\cos\theta)$—Legendre function of argument $-\cos\theta$ and (complex) order ν,
$\nu+\frac{1}{2}=ka\,S_n$ where S_n, for $n=0, 1, 2, \ldots$, is determined from the boundary condition and is described below, and

$$\delta_n\cong\frac{1}{1+\dfrac{\sin 2kh\,C_n}{2\,kh\,C_n}}\cong 1/2 \text{ for } n=0, \cong 1 \text{ for } n=1, 2, 3, \ldots .$$

The individual terms in the above series correspond to the waveguide modes. In the general case, the factor S_n is obtained from the solution of a complicated transcendental equation which involves spherical Bessel functions of large argument and complex order. Certain aspects of this problem have been discussed by Schumann [1] and also the author [4]. For a homogeneous earth and a homogeneous ionosphere, both assumed isotropic, it has been shown that S_n may be approximated by [17]

$$S_n=\left\{1-\left(\frac{\pi n}{kh}\right)^2\frac{1}{4}\left[1+\sqrt{1+4i\,\frac{\Delta kh}{(\pi n)^2}}\right]^2\right\}^{\frac{1}{2}} \tag{2}$$

where

$$\Delta=\frac{1}{N_i}+\frac{1}{N_g}$$

in terms of the refractive indices, N_g and N_i, of the earth and the ionosphere respectively. This equation is valid subject to the condition $|\Delta|kh<<1$. The sign of the radical is chosen in the above equation which makes the real part positive. The values of S_n then are located in the fourth quadrant of the complex plane. When $n=0$, the above simplifies to

$$S_0=\left[1-i\,\frac{\Delta}{kh}\right]^{\frac{1}{2}}. \tag{3}$$

Now, since $|\Delta|kh<<1$, the radical in eq (1) can be expanded for $n>0$ to yield

$$S_n=\left[1-\left(\frac{\pi n}{kh}\right)^2-i\,\frac{2\Delta}{kh}\right]^{\frac{1}{2}} \text{ for } n=1, 2, 3, \ldots . \tag{4}$$

Furthermore, if in addition,

$$\frac{|\Delta|}{kh}<<1-\left(\frac{\pi n}{kh}\right)^2$$

it is permissible to write, for all n,

$$S_n \cong \left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{\frac{1}{2}} - i\,\frac{\epsilon_n}{2kh}\,\Delta\left[1-\left(\frac{\pi n}{kh}\right)^2\right]^{-\frac{1}{2}} \tag{5}$$

where $\epsilon_0=1$, $\epsilon_n=2$ (for $n=1, 2, 3, \ldots$). This latter equation is valid when the "waveguide modes" are not near cutoff and is in the form given by Schumann [2]. It is quite analogous to the standard results for the propagation constant in a rectangular waveguide with finitely conducting walls [18, 19].

3. Behavior of the Modes for the Spherical Earth

For purposes of computation of the mode series, several simplifications can be made. The asymptotic expansion for the Legendre function, given by

$$P_\nu(-\cos\theta) \cong \left(\frac{2}{\pi\nu\sin\theta}\right)^{\frac{1}{2}} \cos\left[\left(\nu+\frac{1}{2}\right)(\pi-\theta)-\frac{\pi}{4}\right] \tag{6}$$

is valid if $|\nu| >> 1$ and θ not near 0 or π [20]. Thus, the modes are simply proportional to

$$\frac{1}{(\sin\theta)^{\frac{1}{2}}} \cos\left[kaS_n(\pi-\theta)-\frac{\pi}{4}\right]$$

which apart from a constant factor can be identified as the linear combination of two peripheral waves. These have the form

$$\frac{1}{(\sin\theta)^{\frac{1}{2}}}\, e^{-ikaS_n\theta}$$

and

$$\frac{1}{(\sin\theta)^{\frac{1}{2}}}\, e^{-ikaS_n(2\pi-\theta)} e^{i\pi/2}$$

where $\theta < \pi$.

These waves are traveling in opposing directions along the two respective great circle paths $a\theta$ and $a(2\pi-\theta)$, from the source to the observer. It is noticed that there is a $\pi/2$ phase advance which the wave traveling on the long great circle path picks up as it goes through the pole $\theta=\pi$. The linear combination of these two traveling waves is to form a standing wave pattern whose distance δ_m between minimums is approximately given by

$$k\delta_m \operatorname{Re} S_n = \pi \text{ or } \delta_m = \lambda/(2 \operatorname{Re} S_n)$$

subject to

$$\operatorname{Im} S_n \ll \operatorname{Re} S_n.$$

The preceding asymptotic expansion for $P_\nu(-\cos\theta)$ is not usable at and in the vicinity of the pole $\theta=\pi$. In this region a suitable representation is given by [20]

$$P_\nu(-\cos\theta) = J_0(\eta) + \sin^2\left(\frac{\pi-\theta}{2}\right)\left[\frac{J_1(\eta)}{2\eta} - J_2(\eta) + \frac{\eta}{6} J_3(\eta)\right] + 0\left(\sin^4\left(\frac{\pi-\theta}{2}\right)\right) \tag{7}$$

where $\eta=(2\nu+1)\sin((\pi-\theta)/2)$. $J_m(\eta)$, for $m=0, 1, 2,$ and 3, is the Bessel function of first type of argument η and order m. When $\pi-\theta$ is small, the first term is usually sufficient and furthermore

$$\eta \cong \left(\nu+\frac{1}{2}\right)(\pi-\theta) \cong kaS_n(\pi-\theta).$$

Thus for mode n, the field in the neighborhood of the pole is proportional to the Bessel function

$$J_0[kaS_n(\pi-\theta)].$$

It is then not surprising to see that the first term of the asymptotic expansion of J_0 is the same as that of P_ν $(-\cos\theta)$.

At the lower end of the ELF band, where the wavelength is large compared to the height of the ionospheric reflecting layer, the electric field is essentially radial and only one waveguide-type mode is significant. The field is thus expressed by the first term of the mode series which reads

$$E_r=\frac{I\,ds\,\eta}{4kha^2}\,\frac{\nu(\nu+1)}{\sin\nu\pi}\,P_\nu(-\cos\theta) \tag{8}$$

where $\nu+\frac{1}{2}\cong kaS_0$ and

$$S_0\cong 1-\frac{i}{2kh}\left(\frac{1}{N_i}+\frac{1}{N_g}\right)$$

in terms of the relative refractive indices N_i and N_g of the homogeneous ionosphere and the homogeneous ground, respectively. Now as mentioned above, the factor $P_\nu(-\cos\theta)$ may be replaced by an asymptotic expansion if $ka\theta$ or $ka(\pi-\theta)$ is somewhat greater than unity. The field in this case may be regarded as two azimuthal-type traveling waves. Furthermore, at the pole (θ near π) where the second of these restrictions is violated, it is possible to use an equivalent representation which correctly accounts for the axial focusing. An alternate viewpoint which is suitable at ELF is to consider the field as a superposition of cavity-resonator type modes. It is expected that such a representation would be very good when the circumference of the earth is becoming comparable to the wavelength. A suggestion of this kind was apparently first put forth by Schumann.

The starting point is the expansion formula

$$\frac{P_\nu(-x)}{\sin\nu\pi}=-\frac{1}{\pi}\sum_{n=0}^{\infty}P_n(x)\,\frac{2n+1}{n(n+1)-\nu(\nu+1)} \tag{9}$$

where the summation is over integral values of n. This result follows directly from a formula given by Magnus and Oberhettinger [20] (p. 57) which is valid for $\nu\neq 0,\ \pm 1,\ \pm 2,\ \ldots$, and $0\leqq\theta<\pi$. The electric field, for $h/a<<1$, is thus written

$$E_r=\frac{I\,ds\,\nu(\nu+1)}{4\pi a^2\epsilon_0\omega h}\sum_{n=0}^{\infty}P_n(x)\,\frac{2n+1}{n(n+1)-\nu(\nu+1)} \tag{10}$$

where $x=\cos\theta$. The early terms of the series are then proportional to

$$\begin{aligned}P_0(x)&=1,\\ P_1(x)&=\cos\theta,\\ P_2(x)&=\tfrac{1}{2}(3\cos^2\theta-1),\end{aligned}$$

and so on.

Retaining just the first term it is seen that

$$E_r^\circ=E_r]_{n=0}=\frac{I\,ds}{4\pi a^2\epsilon_0 h}\,\frac{1}{i\omega} \tag{11}$$

which is independent of θ. Clearly this corresponds to a concentric spherical capacitor energized by a current Ids/h resulting in a constant voltage hE_r° between the plates. On rewriting eq (11) in the form

$$hE^\circ=\frac{(I\,ds/h)}{i\omega C_e}$$

it is seen that

$$C_e = \frac{4\pi a^2 \epsilon_0}{h}$$

which can be identified as the capacity between the spherical surfaces whose areas are both $4\pi a^2$ within the approximation $h/a << 1$.

4. Earth-Flattening Approximation

The possibility that the curvature of the earth may be neglected is now investigated. To simplify field calculations at short distances, it would seem desirable to transform the mode series to a form where the first term corresponds to the model of a flat earth and the succeeding terms are corrections for curvature. This approach has been used by Pekeris [21] and more recently by Koo and Katzin [22] in their investigations of microwave duct propagation. From their work, it may be shown that

$$\frac{P_\nu(-\cos\theta)}{\sin \nu\pi} \cong H_0^{(2)}(kS_n\rho) - \frac{1}{12}\left(\frac{\rho}{a}\right)^2 H_2^{(2)}(kS_n\rho) + \text{terms in } (\rho/a)^4, (\rho/a)^6, \text{ etc.} \tag{12}$$

where $\nu + 1/2 = kS_n\rho$ and $\rho = a\theta$. $H_0^{(2)}$ and $H_2^{(2)}$ are Hankel functions of the second kind of order zero and two, respectively. When the great circle distance ρ is reasonably small compared to the earth's radius a, only the first term in the expansion need be retained.

For convenience in what follows, it is convenient to express the field components as a ratio to the quantity

$$E_0 = i(\eta/\lambda) Ids(e^{-ik\rho})/\rho \qquad (\eta \cong 120\pi);$$

E_0 is the radiation field of the source at a distance ρ on a perfectly conducting ground. Thus, for both the source and the observer near the ground, it is not difficult to show that

$$E_z = WE_0$$

where

$$W \cong -i\pi \frac{\rho}{h} e^{ik\rho} \sum_{n=0}^{\infty} \delta_n S_n^2 H_0^{(2)}(kS_n\rho), \tag{13}$$

$$E_\rho = -SE_0$$

where

$$S \cong \frac{\pi}{N_g} \frac{\rho}{h} e^{ik\rho} \sum_{n=0}^{\infty} \delta_n S_n H_1^{(2)}(kS_n\rho) \tag{14}$$

and

$$H_\phi = -TE_0/\eta$$

where

$$T \cong -\pi \frac{\rho}{h} e^{ik\rho} \sum_{n=0}^{\infty} \delta_n S_n H_1^{(2)}(kS_n\rho) = -N_g S \tag{15}$$

where terms containing $(\rho/a)^2$, $(\rho/a)^4$, etc., have been neglected.

When $k\rho >> 1$, corresponding to the "far-zone," the above expressions may be simplified since the Hankel functions may be replaced by the first term of their asymptotic expansion. This leads to the compact result

$$\begin{bmatrix} W \\ S \\ T \end{bmatrix} \cong \frac{(\rho/\lambda)^{\frac{1}{2}}}{(h/\lambda)} e^{i\left[\frac{2\pi\rho}{\lambda} - \frac{\pi}{4}\right]} \sum_{n=0}^{\infty} \delta_n \begin{bmatrix} S_n^{\frac{3}{2}} \\ -S_n^{\frac{1}{2}}/N_g \\ S_n^{\frac{1}{2}} \end{bmatrix} e^{-i2\pi S_n\rho/\lambda} \tag{16}$$

which is valid for $\rho >> \lambda$. As expected, the ratio of W to T for a given mode is S_n which for low order or grazing modes is of the order of unity. The ratio of S to T, quite generally, is $-1/N_g$ which is very small compared to unity; in fact, it vanishes for a perfectly conducting ground as it must.

The excitation by a horizontal electric dipole may be treated in a similar manner. A complete discussion of the theory is given elsewhere [17]. The expression for the vertical field E_z of the horizontal dipole Ids which is located at the origin and oriented in direction $\phi=0$ may be written in the form

$$E_z=SE_0\cos\phi \tag{17}$$

where S and E_0 are defined above. By using the reciprocity theorem, this result can also be deduced directly from eq (15). Comparing the above formula for E_z (for a horizontal dipole) with corresponding formula for E_z, for a vertical dipole, it may be easily shown that

$$\frac{E_z \text{ (for horizontal dipole)}}{E_z \text{ (for verticle dipole)}}\cong\frac{\cos\phi}{N_g}\left(1-\frac{1}{N_g^2}\right)^{\frac{1}{2}}\cong\frac{\cos\phi}{N_g} \tag{18}$$

being valid in the asymptotic or far zone. In terms of the ground conductivity σ_g and dielectric constant ϵ_g,

$$\frac{1}{N_g}=\left(\frac{i\epsilon_0\omega}{\sigma_g+i\epsilon_g\omega}\right)^{\frac{1}{2}}\cong\left(\frac{i\epsilon_0\omega}{\sigma_g}\right)^{\frac{1}{2}}.$$

This is a small quantity.

The above formulas for W, S, and T are valid only if $\rho<<a$. If the first curvature correction term is included, it is a simple matter to show, in the "far zone", that this amounts to multiplying the right-hand side by the factor

$$1+\frac{1}{12}\left(\frac{\rho}{a}\right)^2 \quad \text{or} \quad 1+\frac{\theta^2}{12}.$$

Now, returning to the mode series for a spherical earth and using the far field approximation for the Legendre function, it turns out that W, S, T have precisely the same form as in eq (16) except the factor $(\theta/\sin\theta)^{\frac{1}{2}}$ occurs in the right-hand side. Noting that

$$\left(\frac{\theta}{\sin\theta}\right)^{\frac{1}{2}}\cong1+\frac{\theta^2}{12}+\text{terms in } \theta^4,\ \theta^6, \text{ etc.}$$

it is apparent to a second order that the derived curvature correction is the same as from the Hankel function series.

5. Distance and Frequency Dependence

Certain features of the mode series are best illustrated by a graphical plot of field strength versus distance under various values of the parameters. Since $N_g>>N_i$ it may be safely assumed for ELF that $\Delta=1/N_i$. Then on the assumption that the ionosphere is behaving like a conductor, it is possible to write

$$N_i=\left(1-i\frac{\omega_r}{\omega}\right)^{\frac{1}{2}} \tag{19}$$

where

$$\omega_r=\frac{\omega_0^2}{\nu}=\frac{(\text{plasma frequency})^2}{\text{collision frequency}}.$$

The above relation for N_i is an approximation which is valid when the collision frequency ν between electrons and neutral ions is much less than the angular frequency ω. Since ν is of the order of 10^7 in the lower E region, this is certainly valid at ELF. The (angular) plasma frequency ω_0 is given by

$$\omega_0^2=\frac{e^2N}{\epsilon_0 m}$$

where N is the electron density, e and m are charge and mass of the electron, and ϵ_0 is the dielectric constant of free space. ω_r, which is proportional to the ratio N/ν, is believed to be of the order

of 10^5 in the lower region of the E layer. At least this appears to be the effective value deduced from vlf observations for highly oblique incidence [23]. At ELF it appears that the waves are reflected at higher levels in the ionosphere and the effective value of ω_r is somewhat larger, being possibly of the order of 10^6.

It should be noted in passing that ω_r may be replaced by σ_i/ϵ_0 where σ_i is the effective conductivity of the ionosphere and ϵ_0 is the dielectric constant of free space.

To indicate the variation of the field as a function of distance, the magnitude and phase of W and T are shown in figures 1 to 4 for the typical values: $\omega_r=5\times10^5$ and $h=90$ km. Actually $|W|$ and $|T|$, both divided by the square root of the distance ρ (in kilometers), are plotted since the curves become linear at larger distances. The slope of these linear parts of the curves are proportional to the attenuation rate (in decibels per 1,000 km) as usually defined. It is seen immediately that at the shorter distances, the curves are no longer straight. This immediately indicates that considerable caution should be exercised in computing attenuation rates from spectral analyses of "sferics." For example, it is only when the distance ρ exceeds about one-sixth of a wavelength, is it permissible to assume that log $(|W|/\sqrt{\rho})$ or log $(|E_z|\times\sqrt{\rho})$ vary in a linear manner with distance ρ. Similarly, the phase of W (or E_z) varies in a linear manner only when the distance exceeds about one-half wavelength. Similar remarks apply to the magnetic field variations with distance.

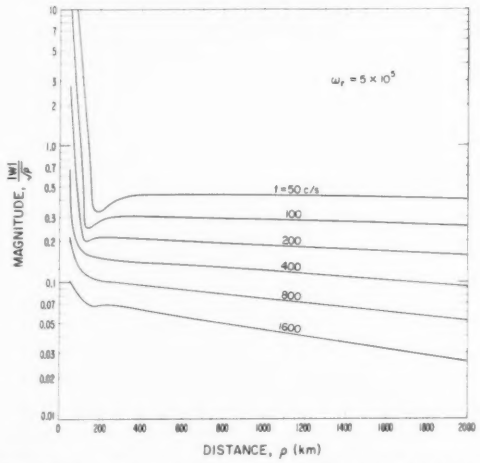


FIGURE 1. *The normalized magnitude of the electric field as a function of distance from the source for* $h=90$ *km.*

(ρ is expressed in kilometers in both ordinate and abscissa).

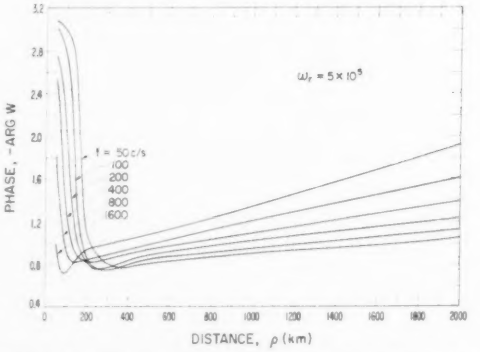


FIGURE 2. *The normalized phase (lag), in radians, of the electric field as a function of distance for* h$=90$ *km.*

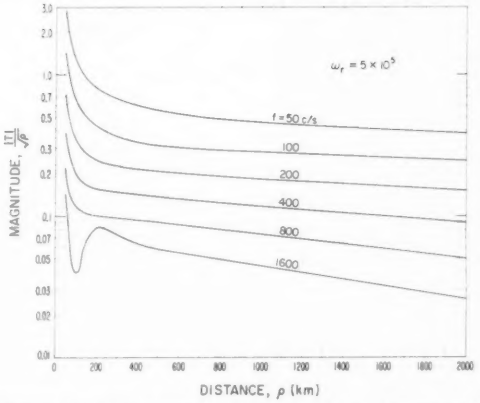


FIGURE 3. *The normalized magnitude of the magnetic field as a function of distance from the source for* h$=90$ *km.*

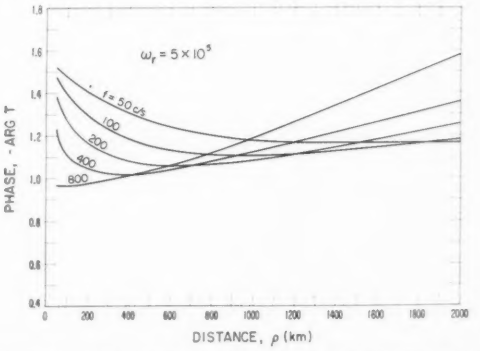


FIGURE 4. *The normalized phase (lag), in radians, of the magnetic field as a function of distance for* h$=90$ *km.*

At the extremely short distances where ρ is somewhat less than the ionospheric reflecting height (i.e., 90 km), the magnitude of W varies essentially as the inverse square of the distance and the phase approaches 180°. In this same region the quantity T varies as the inverse of the distance and the phase approaches 90°. The general characteristics of the field at these extremely short distances are in accord with the near field behavior of a dipole on a flat ground plane when the presence of the ionosphere is neglected. In this rather trivial case

$$W=\left(1-\frac{i}{k\rho}-\frac{1}{k^2\rho^2}\right) \tag{20}$$

and

$$T=\left(1-\frac{i}{k\rho}\right). \tag{21}$$

To shed further light on the nature of the ELF fields in the waveguide, the radial impedance of the wave is now considered. Noting that

$$-\left[\frac{E_z}{\eta H_\phi}\right]_{z=0}=\frac{W}{T} \tag{22}$$

which is by definition the (normalized) impedance of the wave looking in the radial or ρ direction. First it should be noted that for $k\rho>>1$,

$$\frac{W}{T}\cong 1$$

and for $k\rho << 1$ and $\rho << h$,

$$\frac{W}{T}\cong\frac{1}{ik\rho}.$$

For the intermediate and interesting range, $k\rho$ is comparable to unity and ρ is somewhat greater than h. Using the numerical values of W and T mentioned above, the ratio $|W/T|$ and the phase (lag) defined by arg T- arg W are plotted in figures 5 and 6 and as a function of distance for various frequencies. As before, $\omega_r=5\times10^5$ and $h=90$ km.

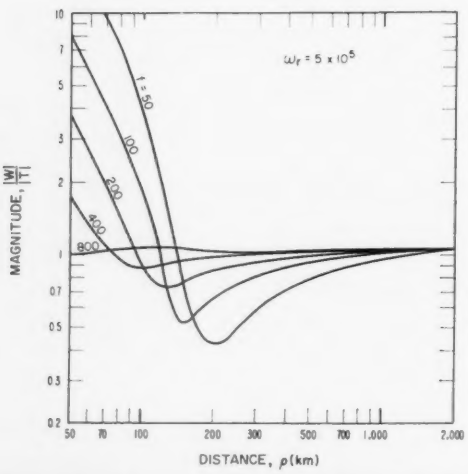


FIGURE 5. *The magnitude of the radial wave impedance.*

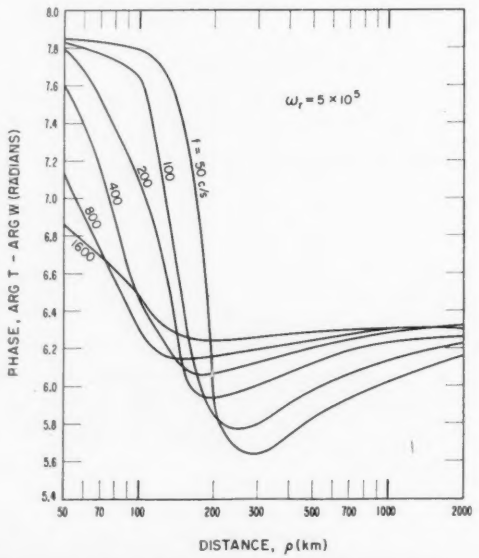


FIGURE 6. *The phase (lag) of the radial wave impedance.*

The impedance ratio W/T is independent of the frequency spectrum of the source provided, of course, it may be represented by an equivalent vertical electric dipole. In fact, this complex ratio could be easily calculated from the frequency spectra of the waveforms of the vertical electric field and the horizontal magnetic field of an atmospheric. The observed variation of magnitude or phase of W/T as a function of frequency should then provide a basis for distance measuring. Such a scheme, while admittedly crude, requires only one receiving station equipped with a vertical whip and a loop antenna.

6. Effect of the Earth's Magnetic Field

In the preceding it has been tacitly assumed that the ionosphere is behaving as an isotropic homogeneous conductor. In this case, the refractive index N_i of the ionosphere may be written [24]

$$N_i \cong \left[1 - i\frac{\omega_r}{\omega}\right]^{\frac{1}{2}} \cong \left(\frac{\omega_r}{i\omega}\right)^{\frac{1}{2}} \tag{23}$$

where

$$\omega_r = \frac{\omega_0^2}{\nu} = \frac{(\text{plasma frequency})^2}{(\text{collision frequency})}.$$

When the earth's magnetic field is steeply dipping, it is appropriate to invoke the quasi-longitudinal approximation of Booker [24]. In this case the refractive index is double valued, one corresponds to the ordinary and the other the extraordinary, thus

$$N_i \cong \left[1 - i\frac{\Omega_r}{\omega}\exp(\pm i\tau)\right]^{\frac{1}{2}} \tag{24}$$

where

$$\tan\tau = \frac{\omega_L}{\nu} = \frac{\text{longitudinal component of gyrofrequency}}{\text{collision frequency}}$$

and

$$\Omega_r = \frac{\omega_0^2}{(\nu^2+\omega_L^2)^{\frac{1}{2}}}.$$

Using this model for the ionosphere, the reflection coefficient for a sharp boundary has been derived by Budden [25]. Adapting this result to ELF it has been shown [17] that the formula for S_n has the same form as the isotropic medium if Δ is defined by

$$\Delta = \frac{1}{N_g} + \left(\frac{i\omega}{\Omega_r}\right)^{\frac{1}{2}} \cos\tau/2. \tag{25}$$

Since $\Omega_r = \omega_r \cos\tau$ it is possible to write the results in the same form as the isotropic ionosphere if we set

$$\Delta \cong \frac{1}{N_g} + \frac{1}{(N_i)_{eff}}$$

where

$$(N_i)_{eff} = \left[\frac{(\omega_r)_{eff}}{i\omega}\right]^{\frac{1}{2}}.$$

$(N_i)_{eff}$ and $(\omega_r)_{eff}$ are the effective values of N_i and ω_r, respectively.

Specifically,

$$(\omega_r)_{eff} \cong \omega_r \frac{\cos\tau}{(\cos\tau/2)^2}. \tag{26}$$

In other words, the effective conductivity of the ionosphere is modified by the factor $\cos \tau/(\cos \tau/2)^2$ which varies from unity to zero as τ varies from 0 to $\pi/2$. Thus, the attenuation is increased as a result of a steeply dipping (or vertical) magnetic field. In fact

$$\frac{\text{attenuation with magnetic field}}{\text{attenuation without magnetic field}} \cong \frac{\cos(\tau/2)}{(\cos \tau)^{\frac{1}{2}}}. \tag{27}$$

If $\tau \ll 1$ (i.e., $\omega_L \ll \nu$), this ratio becomes unity and the influence of the earth's magnetic field vanishes. At the level in the ionosphere where ELF waves are reflected, it is expected that ω and ν are comparable, both being of the order of 10^6. Assuming that they were actually equal, τ becomes 45° and the ratio of the attenuation rates is 1.05. Thus, it is only when ω_L is greater than ν does the earth's magnetic field appreciably influence the attenuation. The ratio is plotted in figure 7 as a function of tan τ or ω_L/ν. This illustrates the situation clearly.

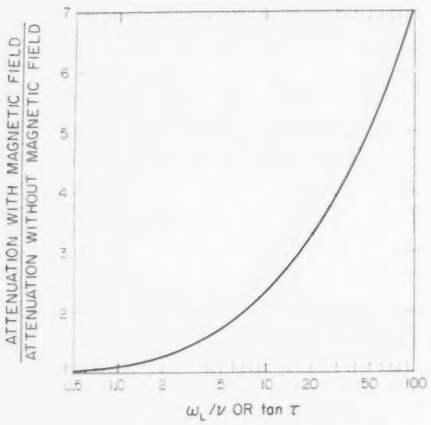


FIGURE 7. *The influence of earth magnetic field on attenuation.*

The abscissa is the ratio of the vertical or longitudinal component of the gyrofrequency, ω_L, and the collision frequency ν.

The preceding results are subject to the validity of Booker's quasi-longitudinal approximation of the Appleton-Hartree equation [24]. The validity of this approximation requires that

$$\frac{\omega_T^4}{4\omega^2\omega_L^2} << \left|\left(1-\frac{\omega_0^2}{\omega^2}-i\frac{\nu}{\omega}\right)^2\right| \tag{28}$$

where ω_L and ω_T are the longitudinal and transverse components of the (angular) gyrofrequency. Clearly this condition is violated when the transverse component of the earth's magnetic field is large such as for propagation along the magnetic equator. This case, however, has been considered by Barber and Crombie [26] who derived explicit results for the reflection coefficient at a sharply bounded ionosphere with a purely transverse magnetic field. Adapting their results to ELF, it is not difficult to show that Δ should now be replaced by [17]

$$\Delta = \frac{1}{N_g} + \left(\frac{i\omega}{\omega_r}\right)^{\frac{1}{2}} X \tag{29}$$

where

$$X = \frac{\left(1+\frac{i\omega}{\omega_r}\right)^{\frac{1}{2}}\left(1+i\frac{\omega}{\omega_r}+i\frac{\omega_T^2\omega}{\nu^2\omega_r}\right)^{\frac{1}{2}} - \frac{\omega_T}{\nu}\left(i\frac{\omega}{\omega_r}\right)^{\frac{1}{2}}}{\left(1+i\frac{\omega}{\omega_r}\right)^2 - \frac{\omega_T^2}{\nu^2}\frac{\omega^2}{\omega_r^2}}.$$

Now at ELF, $\omega \ll \omega_r$ and thus X is very close to unity since ω_T and ν are of the same order of magnitude.[3] Consequently, a transverse magnetic field appears to have a negligible effect on the attenuation and the phase for ELF.[4]

A reasonable conclusion from the above is that only the vertical component of the earth's magnetic field is effective in ELF propagation. Furthermore, the ionosphere is *effectively* an isotropic conductor even in the presence of the earth's magnetic field.

7. Effect of an Inhomogeneous Ionosphere

The electron density in the actual ionosphere usually increases with height in the E region. Furthermore, the collision frequency decreases with height. Thus the effective value of the refractive index cannot be assumed constant. An approach is to let the refractive index increase or decrease from some initial value $\overline{N}$ at height h in a monotonic fashion.[5] Choosing an exponential variation the refractive index as a function of height is explicitly

$$\begin{aligned} N(z) &= 1.0 \quad \text{for } 0<z<h \\ &= \overline{N} \exp\left[(z-h)/l\right] \quad \text{for } z>h \end{aligned} \tag{30}$$

where l is a scale factor. If $l>0$, the refractive index is increasing with height and, if $l<0$, the refractive index is decreasing with height. It has been shown elsewhere [16] that the factor Q has the forms

$$Q=\frac{K_0(i\overline{N}kl)}{K_1(i\overline{N}kl)} \quad \text{if } l>0 \tag{31}$$

and

$$Q=\frac{I_0(-i\overline{N}kl)}{I_1(-i\overline{N}kl)} \quad \text{if } l<0. \tag{32}$$

where I_0, I_1, K_0, and K_1 are modified Bessel functions with their conventional meaning.

The preceding results are valid for $|\overline{N}|>>1$ which is well satisfied at ELF. To the same approximation

$$\overline{N} \cong \left[1-i\frac{\overline{\omega}_r}{\omega}\right]^{\frac{1}{2}} \cong \left(\frac{\overline{\omega}_r}{\omega}\right)^{\frac{1}{2}} e^{-i\pi/4} \tag{33}$$

where $\overline{\omega}_r=\overline{\omega}_0^2/\overline{\nu}$ in terms of the plasma frequency $\overline{\omega}_0$ and collision frequency $\overline{\nu}$. It then follows that the imaginary and real parts of the propagation constant kS_0 are given by

$$\operatorname{Im} kS_0=-\frac{1}{2h}\left(\frac{\omega}{2\omega_r}\right)^{\frac{1}{2}} P(x) \tag{34}$$

and

$$\operatorname{Re} kS_0=k+\frac{1}{2h}\left(\frac{\omega}{2\omega_r}\right)^{\frac{1}{2}} P'(x) \tag{35}$$

where

$$P(x)=|Q| \sin\left(\frac{\pi}{4}-\arg Q\right)\sqrt{2}$$

$$P'(x)=|Q| \cos\left(\frac{\pi}{4}-\arg Q\right)\sqrt{2}$$

---

[3] Note the positive values of ω_T correspond to propagation from east-to-west while negative values correspond to propagation from west-to-east.

[4] If ω_T is somewhat greater than ν, it is seen that to a first order

$$X \cong 1-(\omega_T/\nu)\ (i\omega/\omega_r)^{1/2}$$

which will modify the Re S_0 to a slight extent but not Im S_0.

[5] In this section the influence of the earth's magnetic field is neglected.

and where

$$x=|kl\overline{N}|=|kl|(\overline{\omega}_r/\omega)^{\frac{1}{2}}.$$

The functions $P(x)$ and $P'(x)$ approach unity if x is sufficiently large. This limiting case corresponds to the homogeneous ionosphere. Thus $P(x)$ is the modification of the attenuation and $P'(x)$ is the modification of the phase resulting from a nonhomogeneous ionosphere. It may be observed that the attenuation is generally lower if the refractive index is an increasing function with height. On the other hand, the attenuation is increased where the refractive index is decreasing with height.

To illustrate the behavior of the attenuation rates as a function of frequency, they are plotted in figures 8a and 8b in terms of decibels per 1,000 km of path length for a frequency scale from 50 cps to 1.5 kc. Values of the scale distance l are indicated in kilometers. The height h of the lower edge of the ionosphere is taken to be 90 km. The values chosen for $\overline{\omega}_r$ are 10^5 and 10^6 which are not inconsistent with previous work.

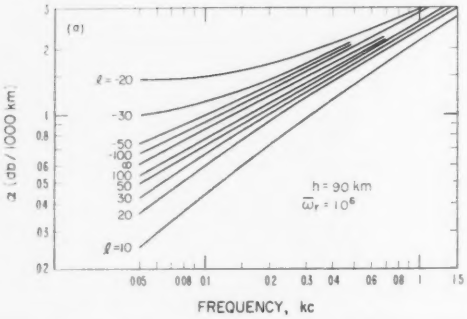


FIGURE 8a. *Attenuation at ELF for an exponential profile,* h=*90 km and* $\overline{\omega}_r=10^6$.

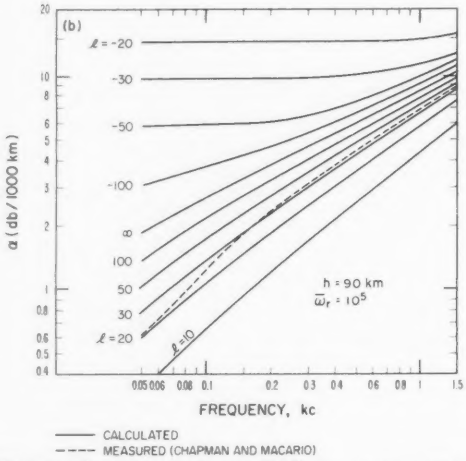


FIGURE 8b. *Attenuation at ELF for an exponential profile,* h=*90 km,* $\overline{\omega}_r=10^5$.

The dashed curve corresponds to nighttime experimental data of Chapman and Macario.

It is seen that the curve for $l=\infty$ corresponding to a homogeneous ionosphere has a slope of 1/2 as it should. Depending on the sign of l and its magnitude, the slope may be modified considerably for the inhomogeneous ionosphere. The dotted curve in figure 8b corresponds to nighttime experimental data from Chapman and Macario [27] who obtained it from the spectral analyses of large numbers of atmospherics recorded in London. Clearly this experimental curve appears to fit fairly well on the l=30-km curve, at least in the range 100 cps to 1.5 kc. There is no reason to expect any better fit than this because of the idealized profile assumed. Furthermore, the atmospherics analyzed by Chapman and Macario were often quite near the receiving location in terms of a wavelength.

8. Propagation of ELF Pulses

In the foregoing discussion, it has been assumed that the source is time harmonic. In most cases of practical interest, the current in the source dipole is of a transient nature such as a surge. The radiated fields are also transient in nature. While the resulting waveforms may be transformed to the frequency plane via spectral analyses, it is often more convenient to study the waveforms themselves [28 to 31].

The source is assumed again to be equivalent to a vertical electric dipole on the earth's surface, but now its current-moment is a function of time and is denoted $p(t)$ which is zero for $t<0$. At a distance ρ along the earth's surface, the resulting vertical electric field is $e_z(t)$ and the horizontal magnetic field is $h_\phi(t)$. The Laplace Transform of the current moment is denoted $P_0(s)$ or more explicitly

$$P_0(s)=Lp(t)=\int_0^\infty e^{-st}p(t)dt \tag{36}$$

where s is the transform variable and may be formally identified with $i\omega$. Also

$$E_z(s)=Le_z(t)$$

and

$$H_\phi(s)=Lh_\phi(t)$$

where L is the Laplace-Transform operator defined above. Now it is assumed that the important frequencies in the spectra are sufficiently low that only the zero-order mode need be retained. Thus

$$\begin{matrix}E_z(s)\\-\eta H_\phi(s)\end{matrix}\cong\frac{\mu s}{2\pi h}P_0(s)\begin{matrix}K_0\\K_1\end{matrix}(sS_0\rho/c) \tag{37}$$

where

$$S_0=\left[1+\frac{c}{h}\left(\frac{\epsilon_0}{s\sigma_i}\right)^{\frac{1}{2}}\right]^{\frac{1}{2}}$$

and where K_0 and K_1 are modified Bessel functions of argument $sS_0\rho/c$. These may be asymptotically represented by

$$K_0(x)\cong\left(\frac{2}{\pi x}\right)^{\frac{1}{2}}e^{-x}\left[1+\frac{A_0}{x}+\frac{B_0}{x^2}+\ldots\right] \tag{38}$$

and

$$K_1(x)\cong\left(\frac{2}{\pi x}\right)^{\frac{1}{2}}e^{-x}\left[1+\frac{A_1}{x}+\frac{B_1}{x^2}+\ldots\right] \tag{39}$$

for large x, where

$$A_0=-1/8,\ B_0=9/128,\ A_1=3/8,\ B_1=-15/128.$$

Thus

$$\begin{matrix}\mathrm{E}_z(s)\\-\eta H_\phi(s)\end{matrix}\cong\frac{\mu}{2\pi h}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(\frac{c}{\rho}\right)^{\frac{1}{2}}s^{\frac{1}{2}}P_0(s)e^{-sS_0\rho/c}\begin{bmatrix}1+\frac{A_0}{s}\frac{c}{\rho}+\frac{B_0}{s^2}\left(\frac{c}{\rho}\right)^2+\ldots\\1+\frac{A_1}{s}\frac{c}{\rho}+\frac{B_1}{s^2}\left(\frac{c}{\rho}\right)^2+\ldots\end{bmatrix} \tag{40}$$

where x has been replaced by $sS_0\rho/c$ in the exponent but is replaced by $s\rho/c$ elsewhere. This is justified since S_0 is near unity. In fact, the binomial expansion

$$S_0=1+\frac{c}{2h}\left(\frac{\epsilon_0}{s\sigma_i}\right)^{\frac{1}{2}}-\frac{1}{8}\left(\frac{c}{h}\right)^2\frac{\epsilon_0}{s\sigma_i}+\frac{1}{16}\left(\frac{c}{h}\right)^3\left(\frac{\epsilon_0}{s\sigma_i}\right)^{3/2}+\ldots \tag{41}$$

may be truncated beyond the second term in most cases of interest, although the third term is retained since it doesn't complicate matters.

The current moment is now taken to be an impulse, that is

$$p(t)=p_0\delta(t)$$

where $\delta(t)$ is the unit impulse function. Thus

$$P_0(s)=p_0.$$

The transforms may now be readily inverted

$$\begin{matrix} e_z^0(t) \\ -\eta h_\phi^0(t) \end{matrix} \cong \frac{\mu}{2\pi^2 h}\left(\frac{c}{2\rho}\right)^{\frac{1}{2}} \frac{\rho_0}{\beta^{3/2}} \exp\left[\frac{\rho c}{8h^2}\left(\frac{\epsilon_0}{\sigma_i}\right)\right]\left[P\left(\frac{t'}{\beta}\right)+\frac{A_0}{A_1}\left(\frac{ct'}{\rho}\right)P_a\left(\frac{t'}{\beta}\right)+\frac{B_0}{B_1}\left(\frac{ct'}{\rho}\right)^2 P_b\left(\frac{t'}{\beta}\right)+\ldots\right]u(t-\rho/c) \tag{42}$$

where

$$P(T)=\left(\frac{1}{2T}-1\right)\left(\frac{1}{T}\right)^{3/2}\exp\left(-\frac{1}{4T}\right) \tag{43}$$

$$P_a(T)=2\left(\frac{1}{T}\right)^{3/2}\exp\left(-\frac{1}{4T}\right) \tag{44}$$

$$P_b(T)=4\left(\frac{1}{T}\right)^{3/2}\exp\left(-\frac{1}{4T}\right)-2\frac{1}{T^2}\pi^{\frac{1}{2}}\operatorname{erfc}\left(\frac{1}{2T^{\frac{1}{2}}}\right) \tag{45}$$

and

$$t'=t-\rho/c \quad \text{and} \quad \beta^{\frac{1}{2}}=\frac{\rho}{2h}\left(\frac{\epsilon_0}{\sigma_i}\right)^{\frac{1}{2}}.$$

It is immediately seen that if $(ct'/\rho)<<1$, the field responses are both proportional to $P(T)$. To illustrate the nature of this function, it is plotted in figure 9 as a function of T. To facilitate the application a multiple scale is included which relates actual time t' in microseconds with the parameter T for distances ρ of 1,000 to 4,000 km. The height of the ionospheric reflecting layer is taken as 70 and 90 km. The quantity B indicated on the curves is related to the ionospheric conductivity by

$$B=\frac{1}{60\sigma_i h}=\frac{2\pi c}{\omega_r h}.$$

The values of B shown, namely 0.05 and 0.1, are typical. For $h=90$ km, these correspond to ω_r values of 4×10^5 and 2×10^5, respectively.

The responses $e_z(t)$ and $h_\phi(t)$ to a general source $p(t)$ can be expressed in terms of the impulse responses $e_z^0(t)$ and $h_\phi^0(t)$, respectively, by using the convolution theorem. Thus

$$e_z(t)=\frac{1}{p_0}\int_{\rho/c}^{t} p(t-\tau)\, e_z^0(\tau)\, d\tau \cdot u(t-\rho/c) \tag{46}$$

and similarly for $h_\phi(t)$. For example, if a special analytical form for $p(t)$ is chosen such as

$$p(t)=\frac{p_0}{2}\left(\frac{6t_m}{\pi}\right)^{\frac{1}{2}}\frac{1}{t^{3/2}}\exp\left(-\frac{3t_m}{2t}\right), \tag{47}$$

the convolution integral may be evaluated to give

$$e_z(t)\cong-\eta h_\phi(t)\cong\frac{\mu}{2\pi^2 h}\left(\frac{c}{2\rho}\right)^{\frac{1}{2}}\frac{p_0}{\gamma^{3/2}}\exp\left[\frac{\rho c}{8h^2}\left(\frac{\epsilon_0}{\sigma}\right)\right]P(t'/\gamma) \tag{48}$$

where

$$\gamma^{\frac{1}{2}}=\beta^{\frac{1}{2}}+(6t_m)^{\frac{1}{2}}=\frac{\rho}{2h}\left(\frac{\epsilon_0}{\sigma_i}\right)^{\frac{1}{2}}+(6t_m)^{\frac{1}{2}}.$$

The function $P(T)$ was defined above and is the same as the one plotted in figure 9. Now, however, T is to be identified with t'/γ rather than t'/β. The special form for $p(t)$ given above is a unidirectional pulse which reaches its maximum value at $t=t_m$ and decays to 7 percent of

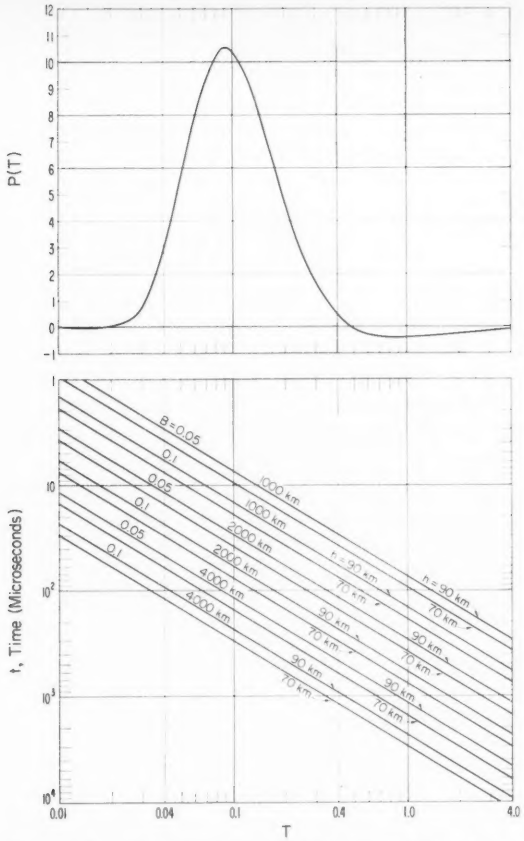


FIGURE 9. *Transient response of the zero-order mode for an impulsive vertical dipole source.*

The lower chart facilitates conversion from the parameter T to actual time t'.

its maximum value at $t=3t_m$. If the duration, $3t_m$, of the source pulse is much less than β the response $e_z(t)$ approaches $e^3(t)$. This is not surprising since $p(t)$ approaches $p_0\delta(t)$ if t_m approaches zero.

The general effect of the finite duration of the source pulse is seen to "stretch" the waveform of the electric and magnetic field. This effect has been discussed at some length in a recent paper by the author entitled "On the theory of the slow-tail portion of atmospheric waveforms." This is to appear in the July issue of the Journal of Geophysical Research.

While most cloud-to-ground lightning strokes may be represented by a vertical electric dipole, it is believed that cloud-to-cloud discharges are better represented by a horizontal electric dipole [32]. In the ELF portion of the spectrum the height of the discharge is very small compared with the wavelength, thus eq (17) is valid. If the source moment is again $p(t)$ then the transform for the vertical electric field, in the far zone, is given by

$$E_z(s)\cong\frac{\mu}{2\pi h}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(\frac{c}{\rho}\right)^{\frac{1}{2}} sP_v(s)e^{-sS_0\rho/c}\left(\frac{\epsilon_0}{\sigma_i}\right)^{\frac{1}{2}}\cos\phi$$

where ϕ is the angle subtended by the horizontal dipole and the direction to the observer. Letting the source be an impulse [i.e., $p(t)=p_0\delta(t)$], the transform may be readily inverted to give

$$e_z(t)\cong\frac{\mu}{2\pi^2\rho}\left(\frac{c}{2\rho}\right)^{\frac{1}{2}}\frac{p_0}{\beta^{3/2}}\left(\frac{\sigma_i}{\sigma_g}\right)^{\frac{1}{2}} Q(T) \tag{49}$$

where

$$Q(T)=\frac{1}{T^{5/2}}\left(\frac{1}{2T}-3\right)\exp\left(-\frac{1}{4T}\right), \tag{50}$$

and where, as before,

$$T=t'/\beta \quad \text{and} \quad \beta^{\frac{1}{2}}=\frac{\rho}{2h}\left(\frac{\epsilon_0}{\sigma_t}\right)^{\frac{1}{2}}.$$

An immediate generalization is a source which may be imagined as an inclined electric dipole. Lightning discharges from cloud-to-ground and cloud-to-cloud would be seldom purely vertical or horizontal. In view of the small dimensions of the discharge paths in terms of a wavelength, it is permissible to replace the inclined channel by superimposed vertical and horizontal electric dipoles. The resulting responses of the radiated field are thus obtained by superposition. For example, if the current moment of the vertical electric dipole is $p_0\delta(t)$ then the current moment of the horizontal electric dipole is $gp_0\delta(t)$ where g is a positive or negative dimensionless number. For lightning strokes with long horizontal sections, g could be large compared to unity [32]. As before, the direction $\phi=0$ corresponds to the direction of the horizontal dipole component.

Invoking the above mentioned principle of superposition, it readily follows that the far zone expression for the inclined dipole is given by

$$e_z(t)\simeq\frac{\mu p_0}{2\pi^2 h}\left(\frac{c}{2\rho}\right)^{\frac{1}{2}}[P(T)+GQ(T)]\beta^{-3/2}$$

where

$$G=g\left(\frac{\sigma_t}{\sigma_g}\right)^{\frac{1}{2}}\frac{h}{\rho}\cos\phi.$$

To indicate the bahavior of the waveform, the quantity

$$S(T)=\frac{P(T)+GQ(T)}{1+|G|}$$

is plotted in figure 10a and 10b for both positive and negative values of G. The quantity $S(T)$ characterizes the waveform of the inclined dipole energized by an impulsive current. The special case $G=0$, such that $S(T)=P(T)$, corresponds to a vertical dipole source, whereas the special case $G=\infty$ corresponds to a purely horizontal dipole source. Such a range of G values can be expected since cloud-to-ground and intra-cloud discharges are primarily vertical, whereas the cloud-to-cloud discharges and air flashes are mainly horizontal and may have either polarity.

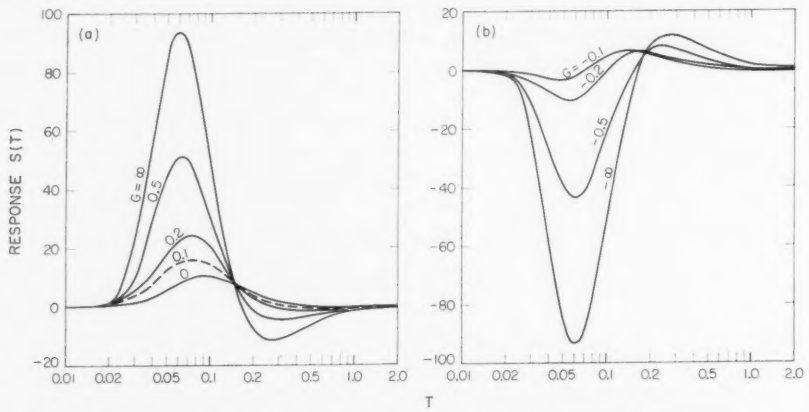


FIGURE 10. *Transient response of the zero-order mode for an inclined dipole source.*
When $G=0$ the dipole is vertical and when $G=\pm\infty$ the dipole is horizontal.

The modification of the pulse shape of the ELF waveforms, as a result of the inclination of the current channel, would appear to be an important factor in interpreting observed data. For example, Pierce [9], Lieberman [29], Tepley [31], and Jean [33] have all observed ELF waveforms which are strikingly similar to those shown in figures 10a and 10b. For example, the observed reception of a pulse which has a second half-cycle of the same magnitude as the first half-cycle can only be reconciled with an inclined source with G equal to about ± 0.2. For certain values of G (e.g., small negative values), a third half-cycle of relatively small amplitude is also produced.

9. Concluding Remarks

In the present theoretical study it has been assumed that the ELF signal propagates entirely in the space between the earth and a concentric ionosphere. It is quite possible that whistler type propagation may also be important. Furthermore, sources of ELF energy may be present in the exosphere due to the possible existence of ionized hydrogen whose gyrofrequency is of the order of 900 cps [34]. Other types of ions such as sodium may also be significant as pointed out by Aarons [14]. Despite these extraneous factors, the evidence that the "slow tails" are propagated in the earth-ionosphere waveguide is overwhelming.

It is the opinion of the author that the only really questionable assumption is the neglect of heavy ions on the constitutive properties of the lower edge of the E region or the top of the D region. Normally, it would be expected that the constitutive properties of the ionosphere would only be influenced by the electrons because of the extremely large ratio between the masses (i.e., of the order of 10^7). However, if the ratio of the number of heavy ions to the number of electrons is of the order of 10^3 or 10^4, as has been suggested by Scott [35] for the Arctic E layer, the influence of heavy ions may be significant at frequencies less than 100 cps. In particular, it should be noted that the gyrofrequency of the sodium ion is approximately 30 cps. However, if the operating frequency is somewhat greater than the gyrofrequency of heavy ions, their influence is expected to be minor.

The author is indebted to Mrs. Nancy Carter and Mrs. Alyce Conda who carried out many of the calculations, to Mrs. Patricia Murdock who typed the manuscript, and to Mrs. Barbara Bolton who prepared the illustrations.

10. References

[1] W. O. Schumann, Über die Ausbreitung sehr langer elektrischer Wellen um die Erde und die Signale der Blitze, Nuovo Cimento **IX** (1952).
[2] W. O. Schumann, Über die Oberfelder bei der Ausbreitung langer, elektrischer Wellen im System Erde-Luft-Ionosphäre und 2 Anwendungen (horizontaler und senkrechter Dipol), Z. angew. Phys. **6,** 34 (1954).
[3] W. O. Schumann, Über die Strahlung langer Wellen des horizontalen Dipols in dem Lufthohlraum zwischen Erde und Ionosphäre Z. angew. Phys. **6,** 225 (1954).
[4] J. R. Wait, On the mode theory of VLF ionospheric propagation, Review Geofisica pura e applicata **37,** 103 (1957) (paper presented at Intern. URSI Conf. Radio Wave Prop. Paris, France, Sept. 1956).
[5] L. Liebermann, Anomalous propagation below 500 c/s, Symp. Prop. VLF Radio Waves, **3,** paper no. 25, Boulder, Colo. (1957).
[6] R. E. Holzer and O. E. Deal, Low audiofrequency electromagnetic signals of natural origin, Nature **177,** 536 (1956).
[7] O. E. Deal, The observation of very low frequency electromagnetic signals of natural origin (unpublished Ph. D. Thesis, Univ. Calif., Los Angeles, Calif., 1956)
[8] F. W. Chapman and W. D. Matthews, Audiofrequency spectrum of atmospherics, Nature **172,** 495 (1953).
[9] F. Hepburn and E. T. Pierce, Atmospherics with very low frequency components, Nature **172,** 837 (1953).
[10] H. F. Willis, Audiofrequency magnetic fluctuations, Nature **161,** 887 (1948).
[11] J. Aarons and M. Henissart, Low frequency noise in the range 0.5 to 20 c/s, Nature **172,** 682 (1953).
[12] R. Beniot, Low frequency radio wave noise of the earth's magnetic field, C. R. Acad. Sci. **242,** 2534 (1956).
[13] P. A. Goldberg, Electromagnetic phenomena of natural origin in the 1.0–150 c/s band, Nature **177,** 1219 (1956).
[14] J. Aarons, Low frequency electromagnetic radiation 10 to 900 c/s, J. Geophys. Research **61,** 647 (1956).
[15] J. M. Watts, An observation of audiofrequency electromagnetic noise during a period of solar disturbance, J. Geophys. Research **62,** 199 (1957).

[16] J. R. Wait, An extension to the mode theory of VLF ionospheric region, J. Geophys. Research **63,** 125 (1958).
[17] J. R. Wait, Terrestrial propagation of VLF radio waves, J. Research NBS **64D,** 153 (1960).
[18] H. Motz, Electromagnetic problems of microwave theory (Methuen & Co., London, 1951).
[19] H. R. L. Lamont, Waveguides (Metheun & Co., London, 1942).
[20] W. Magnus and F. Oberhettinger, Special functions of mathematical physics (Chelsea Publishing Co., New York, N.Y., 1949).
[21] C. L. Pekeris, Accuracy of the earth-flattening approximation in the theory of microwave propagation, Phys. Rev. **70,** 518 (1946).
[22] B. Y.-C. Koo and M. Katzin, An exact earth-flattening procedure in propagation around a sphere, J. Research NBS **64D,** 61 (1960).
[23] J. R. Wait, The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760, (1957).
[24] J. A. Ratcliffe, Magneto-ionic theory (Cambridge Univ. Press, 1959).
[25] K. G. Budden, The reflection of very low frequency radio waves at the surface of a sharply bounded ionosphere with superimposed magnetic field, Phil. Mag. **42,** 504 (1951).
[26] N. F. Barber and D. D. Crombie, V.L.F. reflections from the ionosphere in the presence of a transverse magnetic field, J. Atmospheric and Terrest. Phys. **16,** 37 (1959)
[27] F. W. Chapman and R. C. V. Macario, Propagation of audiofrequency radio waves to great distances, Nature **177,** 930 (1956).
[28] W. O. Schumann, Über die zeitliche Form und das Spektrum ausgesendeter Dipolsignale in einer dielectrischen Hohlkugel mit leitenden Wanden, Verlag Bayerischen Akad. Wiss., Munich (1956).
[29] L. Liebermann, Extremely low frequency electromagnetic waves, II,propagation properties, J. Appl. Phys. **27,** 1477 (1956).
[30] J. R. Wait, Propagation of very-low-frequency pulses to great distances, J. Research NBS **61,** 187 (1958) RP2898.
[31] L. R. Tepley, A comparison of sferics as observed in the VLF and ELF bands, J. Geophys. Research **64,** 2315 (1959).
[32] E. T. Pierce, The development of lightning discharges, Quart. J. Roy. Meteorol. Soc. **81,** 229 (1955). See also, H. Norinder and E. Knudsen, Analysis of daylight photographs of lightning discharges, Recent Advances in Atmospheric Electricity, p. 503 (1958).
[33] A. G. Jean, private communication.
[34] J. W. Dungey, The physics of the ionosphere, p. 229 (The Physical Society, 1955).
[35] J. C. W. Scott, The gyro-frequency in the arctic *E* layer, J. Geophys. Research **56,** 1 (1951).

10.1. Additional References

H. Bremmer, Terrestrial radio waves (Elsevier Publishing Co., Amsterdam, Netherlands, 1949).
W. H. Campbell, A study of micropulsations in the earth's magnetic field, Sci. Rep. No. 1, Nonr 233 (47), Univ. of Calif. Inst. of Geophys., Los Angeles, Calif. (April 1959).
D. D. Crombie, Differences between east-west and west-east propagation of VLF signals over long distances, J. Atmospheric and Terrest. Phys. 12, 110 (1958).
H. J. Duffus, P. W. Nasmyth, J. A. Shand, and C. S. Wright, Sub-audible geomagnetic fluctuations, Nature **181,** 1258 (1958).
R. M. Gallet, A very low frequency emission generated in the earth's exosphere, Proc. IRE 47, 211 (1959).
A. L. Hales, A possible mode of propagation of the "slow" or tail component in atmospherics, Proc. Roy. Soc. A, **193,** 60 (1948).
C. O. Hines, Heavy-ion effects in audio-frequency radio propagation, J. Atmospheric and Terrest. Phys. **11,** 36 (1957). (Shows that whistler ray directions can differ markedly from the direction of the earth's magnetic field when heavy ions are considered).
H. Kaden, Die Reflexions und Schirmwirkung metallischer Hüllen in einer ebenen elektromagnetischen Welle, Sonderdruck Arch Elek Übertragungen, p. 403 (1957).
W. O. Schumann and H. Konig, Über die Beobachtung von "atmospherics" bei geringsten Frequenzen, Naturwissenschaften **41,** 183 (1954).
W. O. Schumann, Über elektrische Eigenschwingungen des Hohlraumes Erde-Luft-Ionosphäre angeregt durch Blitzentladungen, Z. angew. Phys **9,** 373 (1957).
L. R. O. Storey, A method to detect the presence of ionized hydrogen in the outer atmosphere, Can. J. Phys. **34,** 1153 (1956). (Discusses the effect of protons on the propagation of whistlers in the exosphere, collisions are neglected.)
A. D. Watt and E. L. Maxwell, Characteristics of atmospheric noise from 1 to 100 kc, Proc. IRE **45,** 787 (1957).

Additional numerical data are available in, Field strength calculations for ELF radio waves, by J. R. Wait and N. F. Carter, NBS Tech. Note 52 (PB161553) 1960; available from the Government Printing Office, Washington 25, D.C., for 50 cents.

(Paper 64D4–72)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Studies of Natural Electric and Magnetic Fields[1]

G. D. Garland and T. F. Webster

(February 26, 1960)

Simultaneous measurements of short-period natural electric field variations across western Canada are reported. From these it is indicated that the effect of the varying depth to the Precambrian rocks is the dominant factor. Analysis of the simultaneous magnetic and electric measurements gives a resistivity for the Precambrian basement in excess of 30×10^5 ohm meters.

1. Introduction

In considering the possibilities of earth-current communication, a knowledge of natural electric fields in the earth would appear to have significance for two reasons. First, because the natural effects form a background of noise above which an artificial signal must be detected, and secondly, because measurements of natural fields offer a method of deducing the electrical properties, in place, of portions of the earth's crust.

It has been recognized for some time that natural electric currents in the earth, as detected by measurement of the tangential electric field at the surface, bear a relationship to magnetic disturbances. Recently, Hessler and Wescott [1][2] have shown the variation in magnitude of the electric field as a function of the magnetic K-index at College, Alaska. However, the pricise relationship between the surface electric and magnetic disturbance fields at some point on the earth depends on both the distribution of sources and on the variation of conductivity within the earth. It is only if these effects can be separated that useful information on conductivities can be obtained.

An important simplification is possible if the source of disturbance is extended laterally, so that the arrival at the earth's surface may be approximated by a plane wave. It is well known that the tangential electric and magnetic fields, in perpendicular azimuths, say E_x and H_y, bear a simple relationship at the boundary of a uniform conductor. Then-

$$E_x/H_y=(2\sigma T)^{-\frac{1}{2}} \qquad (1)$$

where σ is the conductivity and T is the period of the disturbance. Cagniard [2] has suggested that this approximation is valid for the earth, and that measurements of the fields will lead to a useful effective or apparent conductivity. However, Wait [3] has pointed out that departures from plane-wave conditions can be treated by a series expansion for E_x/H_y. Retaining the first two terms gives

$$E_x/H_y=\eta+\eta/2\gamma^2H_y[\partial^2H_y/\partial y^2-\partial^2H_y/\partial x^2 +2\partial^2H_x/\partial_x\partial y] \qquad (2)$$

where $\gamma=i\sigma\mu\omega-\epsilon\mu\omega^2$ is the intrinsic propagation constant of the earth, and $\eta=i\mu\omega/\gamma$ is its characteristic impedance. Additional terms in the series involve higher inverse powers of γ, where $1/\gamma$ is characteristically some tens of kilometers. The simplified expression of eq (1) will be valid if the second derivatives involved in eq (2) are sufficiently small.

2. Regional Telluric Measurements

It is obviously important to know the extent to which the fields may vary with position on the earth as a result of changing source distance. In order to examine this, and also the effect of variations in earth conductivities, we have compared earth-current disturbances at a number of points across western Canada (fig. 1). The Dominion Observatory maintains an earth-current and magnetic recording station at Meanook, and observations on the electric

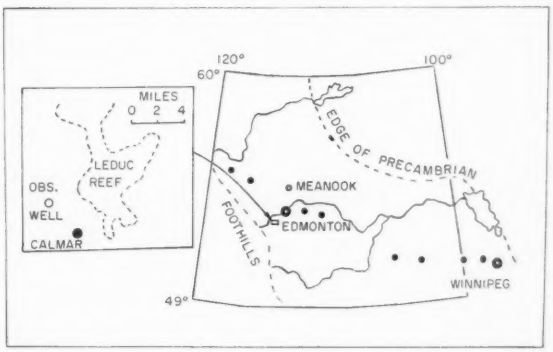


FIGURE 1. *Location of telluric observation points, indicated by solid circles, and the Meanook magnetic observatory.*

Inset, location of well where vertical currents have been detected.

[1] Contribution from Department of Physics, University of Alberta, Edmonton, Alberta, Canada; paper presented at Conference on the Propagation of ELF Radio Waves, Boulder, Colo., Jan. 25, 1960. The experimental work was supported by the National Research Council of Canada.

[2] Figures in brackets indicate the literature references at the end of this paper.

field variations at the other points shown could be compared with simultaneous values at this station. It is the usual practice in earth-current measurements to record the changes in potential across two perpendicular lines (1,000 ft is a convenient length as the disturbances are of the order of millivolts per kilometer), so that any disturbance from the normal potential gradient may be treated vectorially. Within the limit of time resolution possible in these measurements, which were made with a paper speed of 4 in./hr, it was found that events could be readily correlated from all field points to Meanook. The average magnitude of a disturbance, as compared to that recorded at Meanook, varied from station to station, as shown on figure 2. These average magnitudes were obtained by forming the resultant of the two component measurements for a number of disturbances, of period 1 min and greater, observed during several hours of recording at each point. On the same diagram is plotted the depth to the crystalline Precambrian basement, which is formed of rock of much higher resistivity than the overlying sedimentary rocks. There is a fairly good correlation between the two profiles, suggesting that the magnitude of disturbance of the electric field at any point is inversely proportional to basement depth. This would indeed be the case if source effects were completely uniform at all points, and if the basement were completely nonconducting. For then, if h is the basement depth, and σ the uniform conductivity of the sedimentary rocks

$$\sigma hE = I \tag{3}$$

where I is the constant total current flowing in a strip of unit width. At the relatively low frequencies considered here the skin depth is much greater than any basement depth encountered. It was concluded from these results that the major reason for the electric field variation with location was change in effective conductivity, and that source distance effects were minor. The approach employed here is an application, on a very large scale, of a method of geophysical exploration first described by Migaux [4] as the telluric method. It should be noted from figure 1 that the stations were located very nearly

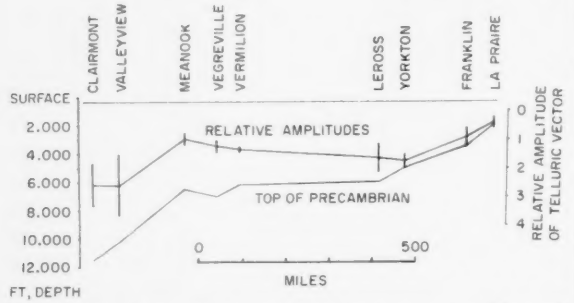


FIGURE 2. *Ratio of telluric disturbance vector at Meanook to that at other stations, and depth of Precambrian basement.*
Vertical lines indicate standard deviations.

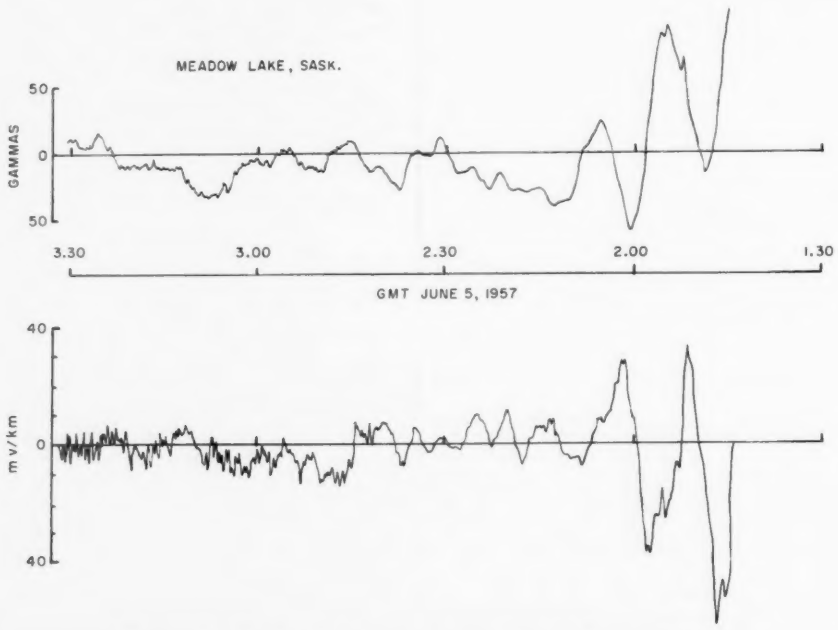


FIGURE 3. *Portion of simultaneous magnetic and telluric records.*
Note the onset of higher frequency events in both records at about 2.45 G.m.t.

along a line of constant magnetic latitude, and different results could well be expected for points at different distances from the auroral zone.

3. Magneto-Telluric Investigations

The apparent uniformity of the natural electric field suggested that Cagniard's approach to determine conductivity at various depths could be useful. Simultaneous measurements, as in figure 3, of electric and magnetic field variations were made at points where the depth to the resistant Precambrian rocks was known. At the time these measurements were made, a flux-gate magnetometer with a limit of sensitivity of 1 gamma (10^{-5} oersted) was used, and all recordings were made with pen recorders. The principle involved in Cagniard's analysis is the extraction of a value of apparent conductivity, by eq (1), for a series of disturbances of different periods. Because of the variation in depth of penetration with period, the apparent conductivity determined will be a function of period, and the form of this functional relationship will depend on the variation of conductivity with depth. Theoretical curves of apparent conductivity against period have been computed for cases of simple layering by Cagniard. It is apparent that the greatest labor involved is in the analysis of the records into discrete frequency components.

In future observations, it is intended that this analysis will be made by digital recording and machine computation. However, the values of apparent resistivity plotted logarithmically against period in figure 4 were obtained by Fourier analysis of the visual records. On the diagram are plotted also two theoretical curves for a two-layer structure. The upper layer is taken to be the sedimentary section, whose thickness is about 1.5 km at this location, with a resistivity, estimated from well logs, of 30 ohm-m. Since the scales are logarithmic, small errors in these assumptions affect the position, but not the form, of the curves. The two theoretical curves are plotted for basement resistivities of 30×10^4 ohm-m and infinity, respectively. While there is considerable scatter in the observed points, the general increase of apparent resistivity with period is indicated. In particular, there is a good indication that the basement resistivity is at least 30×10^5 ohm-m. The observed points indicate a rate of increase of apparent resistivity greater than the theoretical curves, which suggests that the two-layer case is an oversimplification. However, it does appear that an analysis of magnetic and electric fields of this type, especially with more adequate separation of the frequency components, will provide useful information on crystal conductivities. Niblett [5] has extended the analysis to longer periods, using records from the Meanook observatory, and has detected an increase in conductivity at the base of the crust.

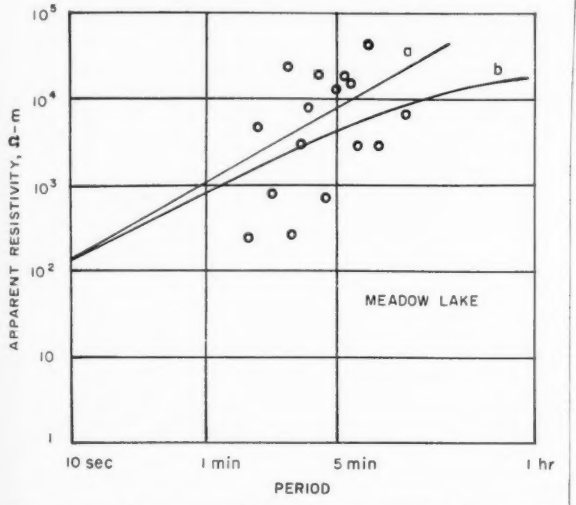


FIGURE 4. *Apparent resistivity by magneto-telluric analysis as a function of period.*

The curves (a) and (b) are the theoretical two-layer variations for basement resistivities of infinity and 30×10^4 ohmmeters, respectively.

4. Vertical Earth Currents

There is some interest in observing the vertical component of electric field within the earth, because this will indicate the amount of deflection of the horizontal currents by local variations in conductivity, and also because of possible correlations with atmospheric measurements. Electrodes at depths of 800 and 3,800 ft have been placed in the well shown in figure 1, and a record of the potential gradient changes is available for a period of about 1 yr. Disturbances in the vertical component amounting to a few percent of the horizontal electric field have been measured, but the ratio of vertical to horizontal field shows a strong dependence on the azimuth of the horizontal current flow. At this particular location, the vertical component is greater when the disturbance current flows north-south, as indicated by the plot of figure 5. This strongly

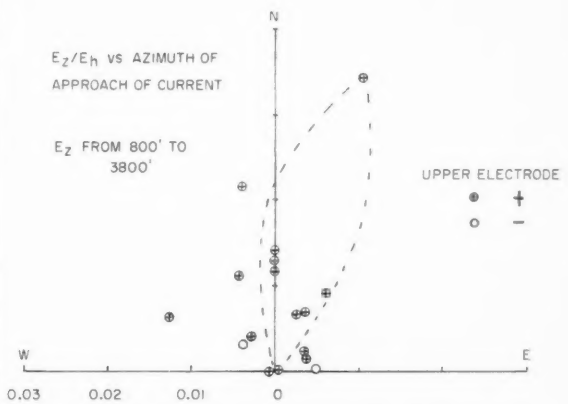


FIGURE 5. *Ratio, obtained from well observations, of vertical to horizontal electric field, plotted radially, as a function of azimuth of horizontal flow, for a series of pulsations from the northerly half-circle.*

suggests lensing of the horizontal currents by a body of anomalous conductivity north or south of the observation well. However, a portion of the observed vertical electric field component in the earth may be associated with the earth-atmosphere current. As the normal component of current density is continuous across the surface, values of 6×10^{13} ohm-m and 30 ohm-m, respectively, for air and earth resistivities indicate a gradient in the earth only 5×10^{-13} times the atmospheric gradient. During a period of atmospheric disturbance, however, the potential gradient in the air may be many times the fair-weather value of 100 v/m, and the corresponding effect in the earth may be detectable. It is hoped that simultaneous measurements in three components, in the earth, and measurements in the atmosphere at this site will indicate the manner in which the air-earth current flows into, or out of, the earth.

5. References

[1] V. P. Hessler and E. M. Wescott, Correlation between earth-current and geomagnetic disturbance. Nature **184,** 627 (1959).

[2] L. Cagniard, Basic theory of the magneto-telluric method of geophysical prospecting, Geophysics **18,** 605 (1953).

[3] J. R. Wait, On the relation between telluric currents and the earth's magnetic field, Geophysics **19,** 281 (1954).

[4] L. Migaux, Une methode nouvelle de géophysique appliquée, Ann. géophys. **2,** 131 (1946).

[5] E. R. Niblett, Variation of electrical conductivity with depth by the magneto-telluric method (abst.), Geophysics **24,** 1150 (1959).

(Paper 64D4–73)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Natural Electromagnetic Energy Below the ELF Range[1]

Wallace H. Campbell

(February 23, 1960)

The transition of natural signals from sferics slow tails to geomagnetic micropulsations was observed between 2.0 and 0.2 cycles per second. Micropulsations with periods of 5 to 30 seconds have characteristics which closely relate to solar terrestrial disturbance phenomena. The low latitude diurnal amplitude variation has maximums at 0945 and 1000 l.m.t. Similar groups of oscillations appear in Alaska and California. Simultaneous pulsation of λ 3914 aurora and magnetic field micropulsations has been observed in Alaska.

1. Introduction

The purpose of this paper is to review some studies of the characteristics of natural electromagnetic fields which are encountered as measurements are made to frequencies lower than those dominated by the sferics slow tails. In particular, magnetic field oscillations with periods of 5 to 30 sec will be discussed. The greater part of this paper is a digest of experimental results which have been previously published. New developments in the relation of magnetic field micropulsations to the pulsating aurora will be reported.

Only the author's work will be discussed in this paper. However, it should be mentioned that recent comprehensive studies of the short period pulsating magnetic field have been made by Holmberg [1],[2] Angenheister [2], Troyičkaya [3], Kato and Watanabe [4], and Shand, Wright, and Duffus [5].

2. Methods

For the field strength observations in frequencies below 20 cps the antennas were large loops, 2 m in diameter with 21,586 turns [6]. The loop natural resonance was at approximately 150 cps. In operation the relatively flat system response was down by 3 db at 0.4 cps and 20 cps for oscilloscope film recordings and at 0.04 cps and 0.4 cps for slow run Esterline Angus pen recorders. Wind velocity and WWV timing were also recorded on the film [6, 7]. The maximum noise level was equivalent to 0.02 gamma and decreased with increasing frequency. Middle latitude observations were made at a southern California site (33°21.5′ N, 116°17′ W) and high latitude experiments were carried out in Alaska (64°42′ N, 148°29.5′ W).

3. Observations

In California the transition of the dominate received natural signals from slow tail sferics to geomagnetic micropulsations was found to occur between 2.0 and 0.2 cps [7]. On a few rare occasions dispersive train oscillations with periods of approximately 1.5 cps and amplitudes reaching 0.015 gamma were observed [6]. Sixty percent of the 4808 hours sampled between March and September 1958, showed sinusoidal type pulsations appearing in groups or packets of four or six oscillations with periods of 20 to 25 sec during the midday hours and falling to 5 to 8 sec by midnight [6]. The periods did not seem to be amplitude dependent. In southern California the signals increased in amplitude at dawn, rose to a maximum of about one-quarter gamma about 0945 to 1000 l.m.t., then decreased to a slight minimum after 1200 hr, showed a secondary maximum at approximately 1400 hr, and faded away gradually in the late afternoon (fig. 1). From comparison with data of researchers [2, 3] in other parts of the world it is apparent that the diurnal pattern of micropulsations is a local mean time variation [7].

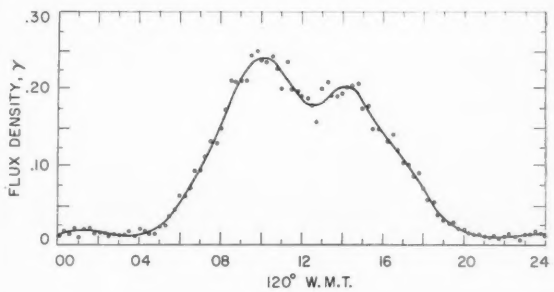


FIGURE 1. *Average diurnal behavior of micropulsations in California from March through September 1958.*

[1] Contribution from Geophysical Institute, University of Alaska; paper presented at Conference on the Propagation of ELF Radio Waves, Boulder, Colo., Jan. 25, 1960.

[2] Figures in brackets indicate the literature references at the end of this paper.

544328—60——9

Average daily amplitudes behave in a manner similar to other phenomena in the solar-terrestrial disturbance family. Good correlation may be obtained with magnetic K and A indices and disturbed F-layer conditions [7]. The 27 day periods of solar activity stands out on the records.

Several occasions a month there appeared "storms" of sudden increases in pulsation amplitude reaching an average of 1.4 gamma and followed for several hours by higher activity than would have been expected for the day [7]. The storm time variation showed a small secondary maximum at 65 min. The occurrence distribution of the 19 storms studied had maximums at 0700 and 1645 u.t. On many occasions the micropulsation storms were coincident with sudden commencement magnetic storms.

At 10:51:30±30 u.t., August 1, 1958, a small group of micropulsations were initiated, lasting for approximately 4 min and attaining an amplitude of approximately 0.2 gamma. About 50 min later a second group of micropulsations commenced, rising gradually to a maximum of 0.3 gamma in 23 min then rapidly disappearing. Since local nighttime activity was rare in the low latitudes and previously accompanied by high magnetic K indices, these oscillations were interpreted as a direct result of the high altitude (70 to 80 km [8]) nuclear explosions at Johnston Island. No micropulsations were observed on the August 12 Johnston Island explosion which was reported to occur at a lower elevation [8].

Twenty-three coincident active groups of micropulsations were observed in 6 days concurrent operation of California and Alaska stations [9]. Although the oscillation periods were similar, peak for peak correspondence was not evident. The diurnal pattern of micropulsations in the auroral zone is dominated by a post midnight maximum typical of the aurora itself [6]. Large nighttime activity in Alaska gave oscillations 10 to 20 times larger than California. Daytime activities were similar at the two stations. In Alaska there seems to be a close relationship between the pulsating aurora and magnetic field micropulsations [10]. Using a lens, interference filter, photomultiplier, and appropriate amplifiers, observations were made of pulsations of the λ3914 and λ5577 auroral lines. Figure 2 shows typical records of corresponding magnetic field micropulsations and pulsating aurora. Figure 3 illustrates the average diurnal behavior of magnetic field micropulsations and pulsating aurora for occasions of joint operation in the month of December 1959. Figure 4 is an illustration of oscillations of the two phenomena in the predawn hours with periods of approximately 25 sec. The optical and magnetic measurements were made at sites separated by 30 miles.

4. Concluding Remarks

Magnetic field micropulsation, though certainly dependent upon solar terrestrial phenomena for their excitation, show many local diurnal features. There are strong indications that the ionosphere is intimately involved in the generation mechanism. It appears that the pulsating aurora is closely related if not identical to the micropulsation phenomena.

The low-latitude work was carried out at the Institute of Geophysics of the University of California and the continuing program in Alaska at the Geophysical Institute of the University of Alaska. The research has had the support of the Electronics Branch of the Office of Naval Research. The assistance of B. Nebel in California and of A. Belon, C. Deehr, M. Rees, and G. Romick in Alaska is gratefully acknowledged.

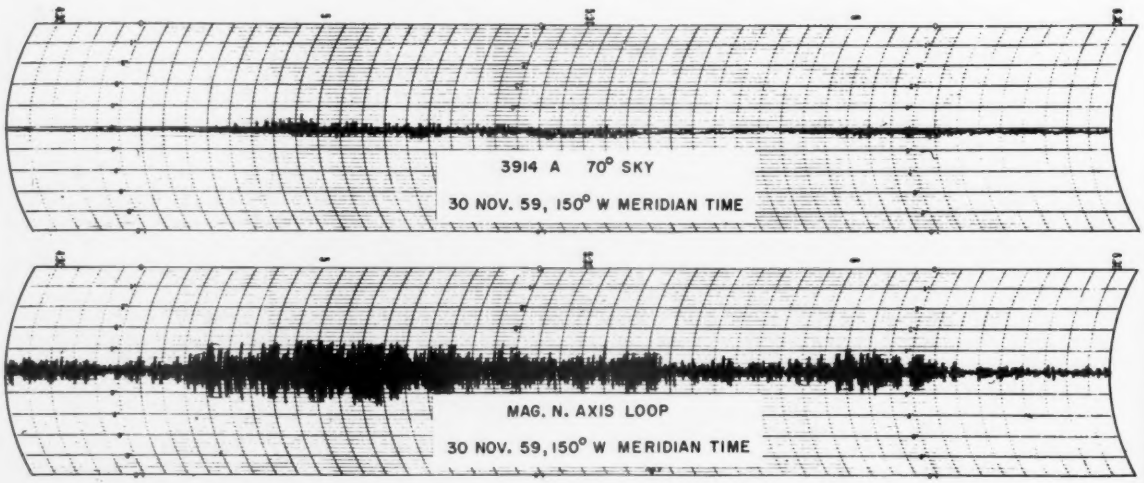


FIGURE 2. *Example of magnetic field micropulsations and pulsation aurora data in Alaska.*

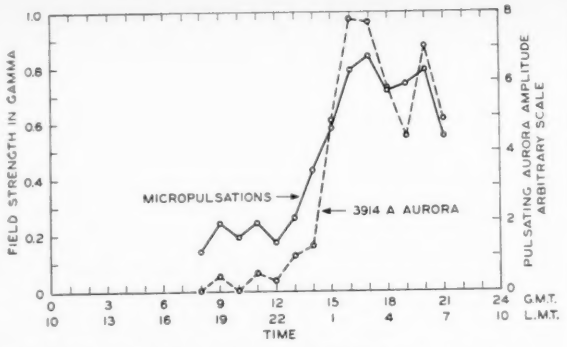


FIGURE 3. *Average diurnal behavior of magnetic field micropulsation and pulsating aurora for the month of December 1959, in Alaska.*

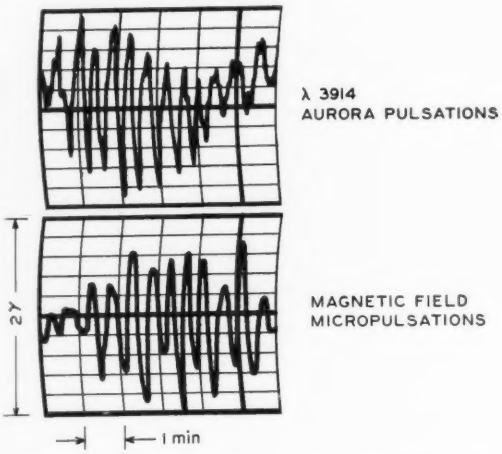


FIGURE 4. *Corresponding pulsations of the aurora and the north magnetic field at stations separated by 30 miles near College, Alaska.*

5. References

[1] E. R. R. Holmberg, Rapid periodic fluctuations in the geomagnetic field, I, Monthly Notices Roy. Astron. Soc., Geophys. Suppl. **6,** 467 (1953).

[2] G. Angenheister, Registrierung erdmagnetischer pulsationen Gottingen, 1952/53, Gerlands Beitr. Geophys. **64,** 108 (1954).

[3] V. A. Troyiçkaya, Earth currents, Priroda **5,** 81 (1955).

[4] Y. Kato and T. Watanabe, A. survey of observational knowledge of the geomagnetic pulsation, Sci. Rept. Tôhoku Imp. Univ., ser 5, Geophys. **8 (3),** 157 (1957).

[5] J. A. Shand, C. S. Wright, and H. J. Duffus, A study of the distribution of geomagnetic micropulsations, Pacific Naval Lab. Rept. 15, D.R.B., Canada, pp. 1 to 29 (1959).

[6] W. H. Campbell, A study of micropulsations in the earth's magnetic field, Inst. Geophys., Univ. Calif. Los Angeles, Sci. Rept. 1, Nonr. 233 (47), 1 to 138 (1959).

[7] W. H. Campbell, Studies of magnetic field micropulsations with periods of 5 to 30 seconds, J. Geophys. Research **64,** 1819 (1959).

[8] S. Matsushita, On artificial geomagnetic and ionospheric storms associated with high-altitude explosions, J. Geophys. Research **64,** 1149 (1959).

[9] W. H. Campbell and B. Nebel, Micropulsation measurements in California and Alaska, Nature **184,** 628 (1959).

[10] W. H. Campbell, Magnetic micropulsations and the pulsating aurora, J. Geophys. Research **65,** 784 (1960).

(Paper 64D4–74)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Possible Application of the System Loss Concept at ELF [1]

Kenneth A. Norton

(February 25, 1960)

A brief description is given of the possible application of the system loss concept at extremely low frequencies. A method for allowing for the effect of external noise levels is outlined.

The system loss concept [1] [2] has been recommended [2, 3] by the Consultative Committee on International Radio as a precise means for describing the overall performance of a radio system including not only the effects of the intervening propagation medium, but also the gains and losses in the transmitting and receiving antennas. The purpose of this brief paper is to call this concept to the attention of engineers designing radio systems which are expected to operate in the extremely low frequency (ELF) or lower frequency ranges. The efficient design of systems operating at such low frequencies is especially difficult, and it is believed that the use of the system loss concept may thus prove to be quite helpful.

Any radio system may be usefully described [4] by means of the simple T network of figure 1. The accessible terminals of the transmitting antenna a and of the receiving antenna b are denoted by AA and BB, respectively. The central member of the T is the mutual impedance $z_m \equiv |z_m| \exp(i\phi)$ between the two antennas, and this contains implicitly all of the effects of the propagation. The self-impedances of the two antennas are z_a and z_b; thus the impedance at the terminals AA, with BB open-circuited, is $z_a \equiv r'_a + ix_a$, while the impedance at the terminals BB, with AA open-circuited, is $z_b \equiv r'_b + ix_b$.

[1] Contribution from Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.; paper presented at Conference on the Propagation of ELF Radio Waves, Boulder, Colo., Jan. 25, 1960.

[2] Figures in brackets indicate the literature references at the end of this paper.

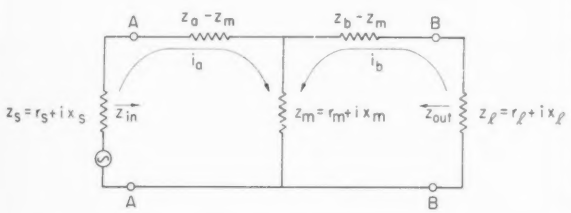


FIGURE 1. *T-network for system loss analysis.*

The system loss is defined as the dimensionless ratio, p_t/p_a, where p_t is the radiofrequency power input to the terminals of the transmitting antenna and p_a is the resultant radiofrequency signal power available at the terminals of the receiving antenna. The system loss is usually expressed in decibels.

$$L_s = 10 \log_{10} (p_t/p_a) = P_t - P_a. \tag{1}$$

In the design of the system it will normally be desirable to minimize L_s. It was shown [1] that this may be accomplished by *simultaneously* matching the source impedance z_s to the complex conjugate of z_{in} and the load impedance $z_l \equiv r_l + ix_l$ to the complex conjugate of z_{out}. For this condition of matching

$$L_s(z_s = z^*_{in};\ z_l = z^*_{out}) = 10 \log_{10} \left[\frac{2r'_a(r'_b + r_l)}{|z^2_m|} - \cos(2\phi) \right] \tag{2}$$

where

$$z_l = r'_b \left\{ 1 - \frac{|z^2_m| \cos(2\phi)}{r'_a r'_b} - \frac{|z^4_m| \sin^2(2\phi)}{4 r'^2_a r'^2_b} \right\}^{\frac{1}{2}} - ix_b + ir'_b \left\{ \frac{|z^2_m| \sin(2\phi)}{2 r'_a r'_b} \right\}. \tag{3}$$

At low frequencies the dominant component of the available noise power p_{an} appearing across the terminals BB will usually have been received after propagation from a noise source with power p_n over propagation paths with an effective system loss $L_n = 10 \log_{10}(p_n/p_{an})$. In this case we see by the following equation that a maximum signal-to-noise ratio may be obtained by minimizing $(L_s - L_n)$.

$$P_a - P_{an} = P_t - P_n - (L_s - L_n). \tag{4}$$

In minimizing $(L_s - L_n)$ it is important to notice that most changes in the receiving system will affect L_s and L_n alike and will thus not improve the signal-to-noise ratio to the extent that the dominant component of the noise is not in the receiving circuits themselves.

It may be helpful to the reader to list some of the advantages and disadvantages of the use of the system loss concept. Its principal advantages are that it: (1) Makes possible a single, precise measurement of the entire system performance; (2) shows explicitly the influence of such circuit parameters as antenna impedances; and (3) combines in a single parameter the interdependent effects of the propagation and the antenna gains and losses.

The principal disadvantage of its use arises from the fact that it does combine the influence of at least partially independent factors. Thus, it is certainly a great advantage to be able to study the transmitting antenna, the propagation medium, and the receiving antenna independently. In those cases where these factors cannot be considered to be independent, the system loss concept must be used for a precise description of the system.

References

[1] K. A. Norton, System loss in radio wave propagation, J. Research NBS, **63D,** 53 (1959).

[2] Documents of the IXth Plenary Assembly of the C.C.I.R., Los Angeles, Calif., Vol. I, Recommendation No. 241, pp. 128 to 130 (1959); and Vol. III, Rept. No. 112, pp. 94 to 99, Intern. Telecommun. Union, Geneva, Switzerland (1959).

[3] K. A. Norton, System loss in radio wave propagation, Proc. IRE **47,** 1661 (1959).

[4] James R. Wait, Transmission of power in radio propagation, Electron. and Radio Engr. **36,** 146 (1959).

(Paper 64D4–75)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Measurements of the Spectrum of Radio Noise From 50 to 100 Cycles Per Second[1]

M. Balser and C. A. Wagner

(January 26, 1960; revised February 19, 1960)

Experimental spectra of radio noise in the band of about 50 to 100 cycles per second have been obtained by means of digital processing. Due to the long integration times which were used, the statistical uncertainty in the estimates of power was reduced to about 3 percent (0.13 decibel). It was hoped in this way to observe maximums in the spectrum due to excitation of higher resonant modes of the earth-ionosphere cavity (for the accuracy of these data, such peaks should be observed if the Q of the cavity were 10 or greater at these frequencies). No statistically significant evidence of these cavity effects was found.

1. Introduction

It has been conjectured for some time that physical effects should be observable due to the resonant modes of the cavity formed by the earth and the D-layer of the ionosphere. The fundamental mode of a cavity of these dimensions has a frequency, f_1 (principally determined by the circumference of the earth rather than the height of the ionosphere) of about 10.6 cps [1].[2] (This estimate depends somewhat on the assumed effective conductivity of the ionosphere, as does the frequency response, or Q of the cavity.) Higher modes obey the relation

$$f_n = f_1\sqrt{\frac{n(n+1)}{2}} \approx \frac{f_1}{\sqrt{2}}\left(n+\frac{1}{2}\right) \quad \text{for } n>3 \text{ or so.} \tag{1}$$

If we assume that the radio noise in this part of the spectrum is the result of excitation of the cavity by sources whose spectrum changes smoothly over intervals of several cycles, then we may expect the response of the cavity to produce bumps in the received noise spectrum, the maximums occurring at the mode frequencies, f_n, and the effective width of the bumps or, alternatively, the peak-to-valley ratio, being determined by the Q of the cavity.

This paper is a description of a preliminary attempt to observe these cavity effects. As will be explained in the next section, this attempt was made in the frequency range above 50 cps, due to equipment limitations, which implies (eq (1)) a measurement of higher modes (at least the seventh), hence at best an indirect attack on the basic problem. Another advantage of operating at these higher frequencies is that the measured fields are still principally due to radiation from sferics. At lower frequencies, probably including the lowest couple of modes, the noise is largely due to motions of local atmospheric charges [2] and to that extent would not exhibit cavity resonances. Despite this disadvantage, the integration technique used here is capable of displaying small power differences which would result from only a part of the noise exhibiting cavity effects, and future work will be attempted at the lower frequencies to avoid the ambiguities arising at the higher ones.

2. Experimental Procedure

It was decided for this first attempt to record some radio noise on tape so that it could then be digitalized and processed on a large scale digital computer (the IBM 709 machine at Lincoln Lab.). Early observations of the low portion of the audio spectrum showed, as might be expected, that most of the power was contained in narrow bands at 60 cps, 180 cps, etc. To avoid this problem, portable equipment was assembled so that the recording could be made far from hum-producing powerlines. Such a spot was found on the Kankamagus Highway between Conway and Lincoln, N.H. One tape was taken on 23 October 1959 (in a light rain) and four more tapes (as well as a test tape of 100 cps) were recorded on a return trip on 11 November 1959.

A block diagram of the equipment is shown in figure 1. The noise signals were received by a 24-ft vertical whip. (The feature which all the cavity modes share, regardless of mode number and direction of the sources, is a vertical electric field, so a vertical whip was chosen as the antenna.) The antenna fed into a low-pass filter which cut off at about 75 kc to prevent the input stages from being overloaded by strong radio signals at HF, resulting in nonlinearities which would put their modulation into the measurement region. Also at the base of the antenna was a cathode follower which fed the signal to a low-pass filter with a 500-cps cutoff and then through 250 ft of cable to the battery-powered audio amplifier. A further filter cut the signal off at about 200 cps so that the signal put on tape at a fixed re-

[1] Contribution from Lincoln Laboratory, a technical center operated by Massachusetts Institute of Technology with the joint support of the Army, Navy, and Air Force; paper presented at Conference on the Propagation of ELF Radio Waves, Boulder, Colo., Jan. 25, 1960.

[2] Figures in brackets indicate the literature references at the end of this paper.

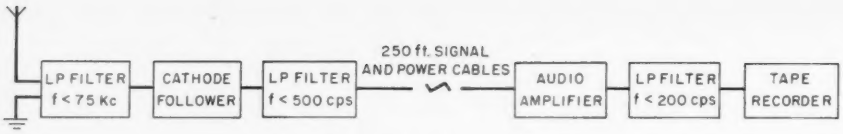


FIGURE 1. *Recording equipment.*

cording level would have a larger power per unit bandwidth. The signal was finally recorded on a battery-operated tape recorder with a spring-driven motor, whose low-frequency response fell off around 50 cps.

The root mean square radio noise voltage in the band below 200 cps was found by comparison with the calibrated 100-cps source to be between ½ and 1 mv. The recording of this same sine wave indicated fluctuations in the tape recorder speed of the order of 1 percent.

3. Digital Processing

Sections of tape somewhat over 10 min long were selected from each of the tapes already mentioned and from a tape of vacuum tube noise recorded by the same system. These were sampled at a rate of 468¾ cps, and the samples were converted to 6-bit numbers (32 levels positive and negative) and recorded on tape which could be read directly into the IBM 709 computer.

A program was written for the 709 to perform the digital equivalent of the analog operations shown in figure 2. The impulse response of the 1-cps filter is taken to be

$$h(t)=\cos \omega t \qquad 0<t<1. \tag{2}$$

FIGURE 2. *Analog of digital processing.*

The output of such a filter for an input signal $x(t)$ is

$$y(t)=\int_{t-1}^{t} x(\tau) \cos \omega(t-\tau)\, d\tau \tag{3}$$

$$=\left(\int_{t-1}^{t} x(\tau) \cos \omega\tau\, d\tau\right) \cos \omega t + \left(\int_{t-1}^{t} x(\tau) \sin \omega\tau\, d\tau\right) \sin \omega t$$

$$\equiv u(t) \cos \omega t + v(t) \sin \omega t$$

where u and v are defined by the previous line. The narrow band noise represented by eq (3) may be rewritten

$$y(t)=A(t) \cos (\omega t+\phi(t)) \tag{4}$$

where the (low-pass) amplitude $A(t)$ is, from (3),

$$A(t)=(u^2(t)+v^2(t))^{1/2}. \tag{5}$$

Finally, the power at $\omega=2\pi f$ is estimated from

$$P=\int_0^T A^2(t)\, dt \tag{6}$$

(constant factors may be neglected since we consider only the power at one frequency relative to another). The program finds u and v from the input tape, $x(t)$, squares and adds them and finally integrates to find P. All the integrals are, of course, taken to be sums over the sample points.

Since $x(t)$ is a sample of (we assume) a stationary random process the derived quantity P is a random variable. Its mean value, $E(P)$, is the quantity we seek. It is shown in the appendix that for gaussian noise the standard deviation, $\sigma(P)$, of this estimate of the power is related to its mean by

$$\frac{\sigma(P)}{E(P)}=\sqrt{\frac{2}{3TW}} \tag{7}$$

where W is the 1-cps bandwidth of the filter and T the integration time. For our parameters, $TW=600$, so $\sigma/E(P)\approx.03(\pm 0.13$ db). (It should be pointed out that the noise was not quite gaussian, due apparently to impulsive peaks. These short peaks, which occurred 3 or 4 percent of the time were clipped so as to allow the normal noise a wider range of quantized values. Analysis shows that this should cause an insignificant change in the spectrum.)

As a test of the digital system (conversion and machine processing) the program was run first on the 100-cps sine wave. The broken curve in figure 3 is the spectrum of the filter defined in eq (2) and centered at 100 cps. (As the reader may observe, the separation between nulls, which is the same as the separation between the frequencies f at which P was estimated is slightly smaller than 1 cps. This was done for programing convenience.) If the input signal were a pure sine wave, the samples P would fall on this curve. As previously noted, the recorded test sine wave varied in frequency. The solid curve in figure 3 shows the response to be expected if the frequency of the input wave were uniformly distributed over a 2-cycle range, or about ± 1 percent.

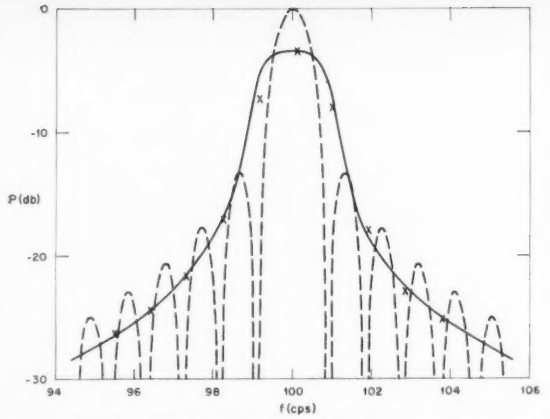


FIGURE 3. *Comparison of computed and theoretical power for test sine wave.*

(This is undoubtedly not precisely the case, but is representative of the averaging behavior of such variations.) The program outputs, P, are shown by crosses. It would appear that the digital system operated properly.

One frequency far removed from 100 cps was also examined to find the noise level of the system. The power at that frequency was some 40 db below that at the frequency of the sine wave. Since there are somewhat over 100 frequency bands, 1 cycle wide, on the tape, and the sine wave was recorded at roughly the same power level as the radio noise, we may expect that the radio noise power is about 20 db over the system noise in any band.

One more statistical consideration deserves mention. If the (effectively) 1-cps filter of figure 2 (the dashed curve of fig. 3) is compared with one centered at the next sampling frequency, separated from the first by 1 cps, it will be seen that there is some overlap in the two filter responses. Thus when a noise waveform is passed through these two filters, they respond to some extent to the same noise components, and the resulting spectrum estimates are therefore not statistically independent. It can be shown that the correlation is proportional to the integral of the product of the two filter spectra [3]. For neighboring frequencies this correlation turns out to be 0.15, actually a rather low correlation. Estimates of the noise spectrum separated by 2 cps or more are, of course, much less correlated.

4. Results

Some representative results of machine computations are shown in figures 4, 5, and 6. The ordinate in each curve is the power P on a linear scale (arbitrary units) as a function of the frequency. The calculated values are at frequencies separated by approximately one cycle and the points are joined by line segments for easier legibility.

The most obvious feature of the spectra in figures 5 and 6 is the peak due to hum pickup at 60 cps (which is actually quite small, amounting to less than 3-db increase in a 1-cps band). Aside from these, there

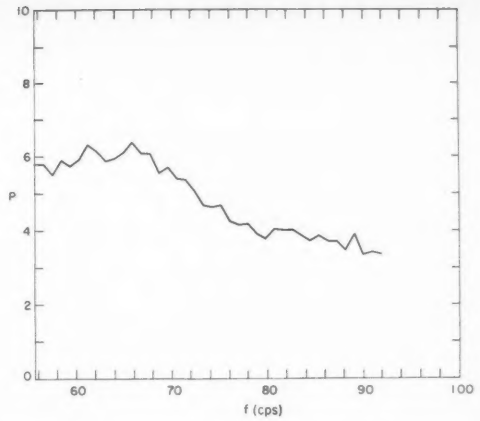


FIGURE 4. *Computed spectrum of vacuum tube noise.*

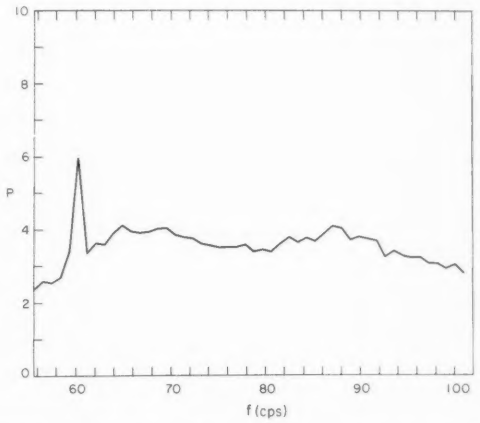


FIGURE 5. *Computed spectrum of tape recorded at 10:45 pm, 11 November 1959.*

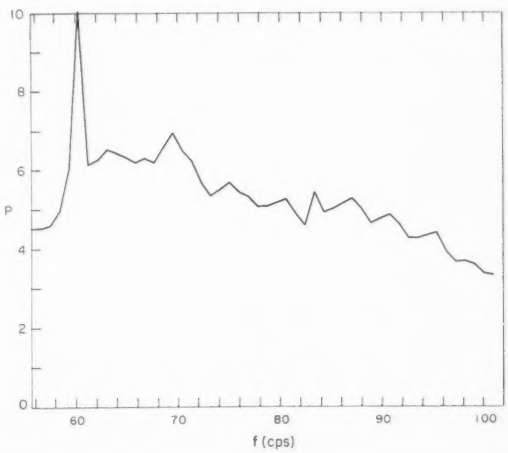


FIGURE 6. *Computed spectrum of tape recorded at 3:30 pm, 11 November 1959.*

are no very pronounced differences in power over distances of a few cycles. The largest peak-to-valley ratio is that near 70 cps in figure 6, which is about 15 percent (0.6 db), while almost all the other fluctuation are of the order of the predicted 3 percent statistical variations (as are, naturally, those of the test noise in fig. 4). We may thus conclude that there is no statistically significant evidence here of cavity effects.

Certain regularities do appear in figure 6 however, which, despite the small size of the fluctuations, attract some attention. Including the one alluded to in the preceding paragraph, the peaks occur at 63, 69, 75, 81, and 87 cps. (With some self-convincing the reader may see equally unsignificant substantiating evidence at 87 cps in figure 5 and in some of the other curves not shown here.) From the expression (eq (1)) for the mode frequencies we find

$$\Delta f = f_{n+1} - f_n \approx \frac{f_1}{\sqrt{2}} \tag{8}$$

and using the 6 cps here indicated for Δf, we derive $f_1 \approx 8.5$ cps, which seems rather low. The frequencies themselves are consistent with the condition that they be multiples of the quantity Δf by integers plus one-half. If we accept the shape of the comparatively pronounced peaks of figure 6 we may, by a computation involving the superposition of identical, periodically spaced resonance curves, compute the (optimistic) value of $Q \approx 10$ for these "modes." This is just about the limit for these calculations, since the peak-to-valley ratio for a set of modes with Q less than about 9 would be smaller than that due to random changes.

Unfortunately, we can accept these figures only as intriguing possibilities or, more probably, as another demonstration of how easy it is to find nonexistent periodicities in spectra. We cannot overemphasize that only one of these peaks is of a size to be considered statistically significant, and the reader may find other "peaks" which do not fall in the conjectured sequence. The conclusion must be, as previously stated, that to the limit of these measurements, no good evidence of cavity effects has been observed. The decisive tests have still to be conducted, possibly with longer integration time, definitely at frequencies nearer, if not at, the fundamental frequency.

5. Appendix

We wish to determine the mean and standard deviation of the estimate, P, of the power of a narrow band gaussian noise

$$y(t) = u(t) \cos \omega t + v(t) \sin \omega t \tag{A1}$$

where

$$P = \int_0^T (u^2(t) + v^2(t)) dt. \tag{A2}$$

u and v are taken to be stationary low-pass gaussian processes, both with mean value $E(u)=0$, mean square $E(u^2)=1$ and autocorrelation $E[u(t)u(t+\tau)] = \rho(\tau)$. Clearly

$$E(P) = \int_0^T [E(u^2) + E(v^2)]\, dt = 2T \tag{A3}$$

$$E(P^2) = \int_0^T \int_0^T [E(u^2(s)u^2(t)) + E(v^2(s)v^2(t)) + E(u^2(s)v^2(t)) + E(u^2(t)v^2(s)]\, ds dt. \tag{A4}$$

Let us assume that the spectrum of $y(t)$ is symmetric about ω, so that $E(u(s)v(t))=0$. Using a standard result about the fourth moment of gaussian noise [4], we obtain

$$E(P^2) = 4T^2 + 4 \int_0^T \int_0^T \rho^2(s-t)\, ds dt. \tag{A5}$$

Thus

$$\sigma^2(P) = E(P^2) - E^2(P) = 4 \int_0^T \int_0^T \rho^2(s-t) ds dt \tag{A6}$$

$$= 8T \int_0^T \left(1 - \frac{r}{T}\right) \rho^2(r)\, dr.$$

For the filter output we are considering, it can be shown either directly from eq (3) or by Fourier-transforming the spectrum in figure 3 that

$$\rho(r) = 1 - |r|W \qquad 0 < |r| < \frac{1}{W}, \tag{A7}$$

where W is the bandwidth of the filter (in this case 1 cps). Since $TW \gg 1$, we may neglect r/T in comparison with unity. Therefore

$$\sigma^2(P) \approx 8T \int_0^{\frac{1}{W}} (1 - rW)^2\, dr = \frac{8T}{3W}, \tag{A8}$$

giving

$$\frac{\sigma(P)}{E(P)} = \sqrt{\frac{2}{3TW}}. \tag{7}$$

We are very grateful to M. Freimer, who wrote the 709 program discussed in section 3. We would also like to thank W. Rutkowski and J. B. Steele, who performed the sampling and digital conversion and recording on IBM tape.

6. References

[1] W. O. Schumann, Über die Dämpfung der elektromagnetischen Eigenschwingungen des Systems Erde-Luft-Ionosphäre, Z. Naturforsch. **7A** 250 (1952).

[2] M. I. Large and T. W. Wormell, Fluctuations in the vertical electric field in the frequency range from 1 cycle per second to 500 cycles per second, Recent Adv. Atmos. Elect. L. G. Smith, Ed., p. 603 (Pergamon Press, New York, N.Y., 1958).

[3] R. B. Blackman and J. W. Tukey, The measurement of power spectra from the point of view of communications engineering, Bell System Tech. J. **37**, Pt. I, p. 185, Pt.II, p. 485 (1958).

[4] W. B. Davenport, Jr. and W. L. Root, An Introduction to the theory of random signals and noise, p. 168 (McGraw-Hill Book Co., Inc., New York, N.Y., 1958).

(Paper 64D4–76)

JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation

Vol. 64D, No. 4, July–August 1960

Listing of Published VLF Symposium Papers

A Symposium on the Propagation of Very-Low-Frequency Radio Waves was held at the Central Radio Propagation Laboratory of the National Bureau of Standards, Boulder, Colo., on January 23 to 25, 1957. At that time, preprints of the papers were available which were supplied to the participants. Since then a large number of requests have been received for additional copies. Since these have not been available for some time, it is felt that the following list of open literature references is of value. These comprise the bulk of the symposium papers in their original or in modified form.

S. A. Bowhill, Ionospheric irregularities causing random fading of very low frequencies, J. Atmospheric and Terrest. Phys. **11,** 91 (1957).

H. Bremmer, Applications of operational calculus to ground-wave propagation, particularly for long waves, IRE Trans. PGAP **AP–6,** 267 (1958).

Kenneth G. Budden, The "waveguide mode" theory of the propagation of very-low-frequency radio waves, Proc. IRE **45,** 772 (June 1957).

J. Chapman and E. T. Pierce, Relations between the character of atmospherics and their place of origin, Proc. IRE **45,** 804 (June 1957).

William Q. Crichlow, Noise investigation at VLF by the National Bureau of Standards, Proc. IRE **45,** 778 (June 1957).

Robert H. Doherty, Pulse sky wave phenomena observed at 100 kc (to be submitted to J. Research NBS, Sec. D).

A. J. Ferraro, J. J. Gibbons, R. E. Houston, W. J. Ross, and E. R. Schmerling, 75 kc/sec pulsed normal incidence ionospheric sounding, (submitted to J. Research NBS, Sec. D).

Forrest F. Fulton, Jr., The effect of receiver bandwidth on the amplitude distribution of VLF atmospheric noise (to be submitted to J. Research NBS, Sec. D).

J. Kenneth Grierson, A technique for the rapid analysis of whistlers, Proc. IRE **45,** 806 (June 1957).

W. E. Gustafson, T. E. Devaney, A. N. Smith, Ground system studies of high power VLF antennas, NEL Rept. 822 (16 December 1957).

Edward L. Hill, Very-low-frequency radiation from lightning strokes, Proc. IRE **45,** 775 (June 1957).

C. O. Hines, Heavy-ion effects in audio-frequency radio propagation, J. Atmospheric and Terrest. Phys. **11,** 36 (1957).

A. G. Jean, L. J. Lange, and J. R. Wait, Ionospheric reflection coefficients at VLF from sferics measurement, Geofis. pura e appl. **38,** 147 (1957).

Martin B. Kraichman, Basic experimental studies of the magnetic field from electromagnetic sources immersed in a semi-infinite conducting medium, J. Research NBS **64D,** 21 (Jan. 1960).

Leonard Liebermann, Extremely low-frequency electromagnetic waves. I. Reception from lightning. II. Propagation properties, J. Appl. Phys. **27,** 1473 (1956).

R. S. Macmillan, W. V. T. Rusch, and R. M. Golden, A very-low-frequency antenna for investigating the ionosphere with horizontally polarized radio waves, J. Research NBS **64D,** 27 (Jan. 1960).

C. A. McKerrow, Some recent measurements of atmospheric noise in Canada, Proc. IRE **45,** 782 (June 1957).

M. G. Morgan and W. C. Johnson, Whistlers and dawn chorus at University of Alaska and their correlation with simultaneous data at the geomagnetic conjugate whistlers and dawn chorus at Knob Hill, Quebec (to be submitted to J. Research NBS, Sec. D).

John A. Pierce, Intercontinental frequency comparison by very-low-frequency radio transmission, Proc. IRE **45,** 793 (June 1957).

H. J. Duffus and J. A. Shand, Some observations of geomagnetic micropulsations, Can. J. Phys. **36,** 508 (1958).

R. L. Smith and R. A. Helliwell, Calculation of the electron density of the outer ionosphere using whistlers (to be submitted to J. Research NBS, Sec. D).

R. L. Smith, The guiding of whistlers in a homogenous medium J. Research NBS **64D** (to be published in Sept. 1960).

Glenn M. Stanley, Layered earth propagation in the vicinity of Point Barrow, Alaska, J. Research NBS **64D,** 95 (Jan. 1960).

L. R. O. Storey, A method for interpreting the dispersion curves of whistlers, Can. J. Phys. **35,** 1107 (1957).

J. R. Wait and A. Murphy, The geometrical optics of VLF sky wave propagation, Proc. IRE **45,** 754 (June 1957).

James R. Wait, The mode theory of VLF ionospheric propagation for finite ground conductivity, Proc. IRE **45,** 760 (June 1957).

James R. Wait, The attenuation vs frequency characteristics of VLF radio waves, Proc IRE **45,** 768 (June 1957).

A. D. Watt and E. L. Maxwell, Characteristics of atmospheric noise from 1 to 100 kc, Proc. IRE **45,** 787 (June 1957).

J. M. Watts, Audio-frequency electromagnetic hiss recorded at Boulder in 1956, Geofis. pura e appl. **37,** 169 (1957).

Arthur H. Waynick, The present state of knowledge concerning the lower ionosphere, Proc. IRE **45,** 741 (June 1957).

Harold A. Wheeler, Fundamental limitations of a small VLF antenna for submarines, IRE Trans. PGAP **AP–6,** 123 (Jan. 1958).

Harold A. Wheeler, Fundamental relations in the design of a VLF transmitting antenna, IRE Trans. PGAP **AP–6,** 120 (Jan. 1958).

H. G. Wolff, High speed frequency shift keying of LF and VLF radio circuits, IRE Trans. **CS–5,** 29 (Dec. 1957).

Irving Yabroff, Reflection at a sharply-bounded ionosphere, Proc. IRE **45,** 750 (June 1957).

Publications of the Staff of the National Bureau of Standards*

Selected Abstracts

The power spectrum of the variation of a carrier envelope in tropospheric scatter propagation, A. F. Barghausen and R. E. McGavin, *Fifth Natl. Commun. Symp. Conv. Record, Oct. 5–7, 1959 (Utica, N.Y., 1959).*

Statistical methods are used in investigating the short time variations in the carrier envelope of a signal received over a tropospheric scatter circuit. Considering samples from seven different paths, comparison of carrier amplitude variations are made for signals of different frequency, under different climatic conditions, over paths of various lengths up to 650 miles, and using antennas of differing apertures.

Utilizing the Fourier transform of the autocovariance function, i.e., the power spectrum, the extent of variation of the carrier envelope can be estimated. It is found that spectra on the same path exhibit no diurnal trend, but that the spectrum at a particular time can show considerable departure from the average spectrum for that path. A definite relationship exists between the maximum fading rate (i.e., the number of positive crossings at some level of the received field where this number is a maximum), and the departure of the individual spectra from a common spectrum. If each spectrum is normalized to the fading rate of the sample, all spectra on the same path approach an average value.

Definite similarity is noted in spectra so normalized but obtained on different paths and in different parts of the world. Averaged normalized spectra from the seven paths show little difference except in slope. However, there is evidence that the slope of the spectrum is a function of θ/Ω, where θ/Ω is a ratio of the angular distance to the antenna beamwidths. It would appear that the power spectrum of a carrier envelope for any arbitrary path can be completely specified by the median field, the maximum fading rate, and the parameter θ/Ω.

In determining the maximum fading rate it is shown that the rate of fading varies inversely with frequency, angular distance, θ, and the parameter θ/Ω.

In addition, a detailed analysis of the percentage of time that the carrier envelope faded below a particular level for a fade duration equal to or greater than predetermined values, was made for four of the paths given above. These studies are useful for modulation studies where it is desired to estimate the percentage message errors, taking into account the carrier fade duration as well as the median carrier to noise ratio.

Propagation and production of electromagnetic waves in a plasma, R. M. Gallet, *Nuovo Cimento Suppl. 13, No. 1, p. 234, (1959).*

A very brief discussion on the non-thermal radiation in a plasma is given. Examples are (1) radio bursts from the planet Jupiter, (2) solar bursts, (3) galactic radio emission, (4) aurorae. The author then goes on to say that the theory of a propagation in a plasma explains the phenomena of "whistlers" and "very-low-frequency emission." Whistlers, which are now well known, are described as trains of waves in the kilocycle range which propagate along the earth's magnetic field with high dispersion and low velocity; vlf emissions are said to result from the mass motion of ionized clouds in the outer ionosphere at 2 or 3 earth's radii. The author then suggests that, using the whistler mode, the electron density and electron temperature of a dense plasma could be measured with microwaves whose frequency is lower than both the plasma and the gyrofrequency.

The remainder of the paper is a short review of the theory of wave propagation in an ionized medium with a superimposed magnetic field.

The Ninth Plenary Assembly of the CCIR, J. W. Herbstreit, *Proc. IRE,* **48,** *45* (1960.)

The International Radio Consultative Committee held its Ninth Plenary Assembly in Los Angeles, Calif., April 2 through 29, 1959. The CCIR studies the technical factors affecting international radio telecommunications and makes recommendations to administrations which are of importance in the international allocation and usage of the radio spectrum. These studies are made by 14 Study Groups, each dealing with certain aspects of international telecommunication. At the Los Angeles meeting, the technical aspects of relatively new techniques for conservation of spectrum space were considered. Transmission and receiving standards, propagation characteristics, monitoring methods, modulation techniques including single sideband and stereo sound broadcasting, and the new international aspects of telecommunication with and between space vehicles were among the many topics considered by the Assembly.

An analysis of time variations in tropospheric refractive index and apparent radio path length, M. C. Thompson, Jr., H. B. Janes, and A. W. Kirkpatrick, *J. Geophys. Research* **65** *193 (1960).*

The National Bureau of Standards has been conducting a series of measurements for a study of the characteristics of the turbulent lower atmosphere and its effect on the accuracy of radio direction-finding, guidance, and geodetic measurement systems. The results of three experiments are presented, one consisting of recordings of refractive index and apparent path length variations at 9400 Mc/s over a 155-mile path on Maui, Hawaii, and the other two consisting of similar measurements made over a 9.5-mile path near Boulder, Colo. The correlation of refractive index and apparent path length fluctuations is discussed as well as the power (variance) density spectra of both variables.

On the diffraction of electromagnetic pulses by curved conducting surfaces, J. R. Wait and A. M. Conda, *Can. J. Phys.* **37,** *1384 (1959).*

Starting with the known steady-state solutions for diffraction by a perfectly conducting convex surface, the corresponding transient responses are derived using Fourier-Laplace inversion. Explicit results are given for an incident wave which varies with time as a step function.

Four physical situations are considered in the present paper. The first two concern the currents excited on a convex metallic surface for vertical and horizontal polarization respectively. The latter two deal with the Fresnel diffraction field of a convex metallic surface in which case the source and observer are at a large distance from the crest on the convex obstacle. Again, this is carried out for both vertical and horizontal polarization.

Atmospheric tides and ionospheric electrodynamics, M. L. White, *J. Geophys. Research* **65,** *153 (1960).*

A brief review is given of the resonance theory of atmospheric tidal oscillations including both thermal and gravitational excitation. Semiempirical wind patterns for various latitudes and the semiempirical variation of solar semidiurnal wind velocities with height are given, based on published data on wind measurements at balloon heights, *E*-region winds, and airglow cell movements. The current work on modern tidal theory is also discussed.

A model of the *F* region above $h_{\max} F2$, J. W. Wright, *J. Geophys. Research* **65,** *185 (1960).*

A simple Chapman model (scale height 100 km) of the ionosphere above the peak of the *F* region is found to be in good agreement with the few observed profiles of this region. The ratio of the above-peak electron content, implied by the model to below-peak electron content is found to vary from about 2.8 to 4.0, also in good agreement with observations. The model is used in conjunction with electron density profiles observed from the ground to construct meridional cross sections along the 75° W geographic meridian.

Radio transmission by ionospheric and tropospheric scatter, *Rept. Joint Tech. Adv. Comm., JTAC. Proc. IRE* **48,** *4 (1960).*

This report on "Radio Transmission by Ionospheric and

Tropospheric Scatter" has been prepared as a supplement to "Radio Spectrum Conservation," an earlier comprehensive report which was completed by the Joint Technical Advisory Committee and published by the McGraw-Hill Book Co., Inc., in 1952. The purpose of "Radio Spectrum Conservation," as stated in the Introduction thereto, was "to analyze and evaluate the current uses of the radio spectrum and to formulate constructive suggestions for the future."

Soon after publication of the report "Radio Spectrum Conservation," two new methods for beyond-the-horizon extended-range radio transmission began to emerge as major advances in the communication art. The importance of these new radio techniques, now usually described as "ionospheric scatter" and "tropospheric scatter," led the Joint Technical Advisory Committee in February 1955, to establish an Ad Hoc Subcommittee on Forward Scatter Transmission. It was the task of this Subcommittee to compile factual data on the new scatter techniques and to prepare a report on "Radio Transmission by Ionospheric and Tropospheric Scatter" to supplement the book "Radio Spectrum Conservation."

In 1955, the Subcommittee, in cooperation with the IRE Professional Group on Antennas and Propagation, sponsored publication of the October 1955, Scatter Propagation Issue of the Proceedings of the IRE. This special issue aided greatly in consolidating and disseminating authoritative information on scatter transmission—and remains a major and lasting contribution to the technical literature.

Recently, the Subcommittee completed preparation of the present report which, hopefully, will serve usefully to augment "Radio Spectrum Conservation" (particularly Chapter 2—Propagation Characteristics of the Radio Spectrum), as well as to summarize those unique features of ionospheric and tropospheric scatter which now must be considered in planning for future efficient utilization of the radio spectrum.

Other NBS Publications

Journal of Research, Vol. 64A, No. 3, May–June 1960. 70 cents.

Phase shift effects in Fabry-Perot interferometry. Charles J. Koester.

Infrared absorption spectrum of methane from 2470 to 3200 cm^{-1}. Earle K. Plyler, Eugene D. Tidwell, and Lamdin R. Blaine.

Elastic constants of synthetic single crystal corundum at room temperature. J. B. Wachtman, Jr., W. E. Tefft, D. G. Lam, Jr., and R. P. Stinchfield.

Radial distribution study of vitreous barium borate. Arthur Bienenstock, Aaron S. Posner, and Stanley Block.

Separation and determination of small quantities of aluminum in steel. Bruce B. Bendigo and Rosemond K. Bell.

Conformations of the pyranoid sugars. II. Infrared absorption spectra of some aldopyranosides. R. Stuart Tipson and Horace S. Isbell.

Journal of Research, Vol. 64B, No. 2, April–June 1960. 75 cents.

Non-self-adjoint boundary value problems in ordinary differential equations. Werner Greub and Werner C. Rheinboldt.

Criteria for the existence and equioscillation of best Tchebycheff approximations. John R. Rice.

Note on the solution of Riccati's differential equation. H. Herbert Howe.

On a generalization of the index notation for absolute tensors of arbitrary order. Edmund H. Brown.

Upper and lower bounds for the center of flexure. Lawrence E. Payne.

Half-round inductive obstacles in rectangular waveguide. D. M. Kerns.

Journal of Research, Vol. 64C, No. 2, April–June 1960. 75 cents.

Measurement of cobalt-60 and cesium-137 gamma rays with a free-air chamber. H. O. Wyckoff.

Apparatus for the measurement of the normal spectral emissivity in the infrared. Arthur G. Maki, Ralph Stair, and Russell G. Johnston.

Electrostatic deflection plates for cathode-ray tubes. I. Design of single-bend deflection plates with parallel entrance sections. II. Deflection defocusing distortion of single-bend deflection plates with parallel entrance sections. Lothar Frenkel.

The functional synthesis of linear plots. J. P. Vinti and R. F. Dressler.

Radiation field from a rectangular source. J. H. Hubbell, R. L. Bach, and J. C. Lamkin.

Microwave attenuation measurements with accuracies from 0.0001 to 0.06 decibel over a range of 0.01 to 50 decibels. G. F. Engen and R. W. Beatty.

Effect of oleophobic films on metal fatigue. H. E. Frankel, J. A. Bennett, and W. L. Holshouser.

Ratio-recording spectroradiometer. Harry K. Hammond III, Warren L. Holford, and Milton L. Kuder.

An intermittent-action camera with absolute time calibration. G. Hefley, R. H. Doherty, and E. L. Berger.

Specification for dry cells and batteries (supersedes Circular 559). NBS Handb. **71** (1959) 25 cents.

Report of the 44th National Conference on Weights and Measures 1959. NBS Misc. Publ. **228** (1959) 65 cents.

Research Highlights of the National Bureau of Standards Annual Report, Fiscal Year 1959. NBS Misc. Publ. **229** (1959) 55 cents.

Basic theorems in matrix theory, M. Marcus. NBS Applied Math. Series **57** (1960) 15 cents.

Survey of Central Radio Propagation Laboratory research in tropospheric propagation 1948–1956, J. W. Herbstreit and P. L. Rice, NBS Tech. Note **26** (PB151385) (1959) $4.00.

An atlas of oblique-incidence ionograms, V. Agy, K. Davies, and R. Salaman, NBS Tech. Note **31** (PB151390) (1959) $2.25.

Distribution of incoming lettermail at the Baltimore, Maryland City Post Office, B. M. Levin and A. E. Newman, NBS Tech. Note **33** (PB151392) (1959) $2.50.

Service area of an airborne television station, M. T. Decker, NBS Tech. Note **35** (PB151394) (1959) 75 cents.

Application of RF micropotentiometers for calibration of signal generators to 1000 Mc, L. F. Behrent, NBS Tech. Note **37** (PB151396) (1960).

Design and construction of a liquid hydrogen temperature refrigeration system, D. B. Chelton, J. W. Dean, and B. W. Birmingham, NBS Tech. Note **38** (PB151397) (1960) 75 cents.

Helium refrigeration and liquefaction using a liquid hydrogen refrigerator for precooling, D. B. Chelton, J. W. Dean, and T. R. Strobridge, NBS Tech. Note **39** (PB151398) (1960).

Mean electron density variations of the quiet ionosphere I - March 1959, J. W. Wright and L. A. Fine, NBS Tech. Note **40–1** (PB151399–1) (1960) $1.25 .

Paramagnetic resonance in the free hydroxyl radical, H. E. Radford, Nuovo Cimento Ser. X **14,** 245 (1959).

Relaxation processes in multistate systems, K. E. Shuler, Phys. of Fluids **2,** No. 4, 442 (1959).

Optical measurements on thin films of condensed gases at low temperatures, J. Kruger and W. J. Ambs, J. Opt. Soc. Am. **44,** 1195 (1959).

The following papers appear in Proc. Intern. Rubber Conf., Nov. 8 to 13, 1959 (Washington, D.C., 1959):

Power loss and operating temperature of tires, R. D. Stiehler, M. N. Steel, G. G. Richey, J. Mandel, and R. H. Hobbs, p. 73.

An indoor tester for measuring tire tread wear, G. G. Richey, J. Mandel, and R. D. Stiehler, p. 104.

Measurement of the aging of rubber vulcanizates, J. Mandel, F. L. Roth, M. N. Steel, and R. D. Stiehler, p. 221.

Standard materials for rubber compounding, F. L. Roth and R. D. Stiehler, p. 232.

Invariant and complete stress functions for general continua, C. Truesdell, Arch. Rat. Mech. Anal. **1,** No. 4, 1 (1959).

On the convergence of Gauss' alternating procedure in the method of least squares, A. M. Ostrowski, Ann. Mat. Pura Appl. (Bologna, Italy) (IV) **48,** 229 (1959).

Wavelength definition of the meter, I. C. Gardner, Systems of Units, National and International Aspects, Publ. Am. Assoc. Advance. Sci. No. **57,** p. 53 (1959).

Ethane carbon-carbon distance obtained from infrared spectra, H. C. Allen and E. K. Plyler, J. Chem. Phys. **31,** No. 4, 1062 (1959).

The anomalous inversion in cristobalite, R. F. Walker, The kinetics of high temperature processes, 228 pages (J. Wiley & Sons, New York, N.Y., 1959).

Digital recording of electrocardiographic data for analysis by a digital computer, L. Taback, E. Marden, H. L. Mason, and H. V. Pipberger, IRE Trans. Med. Electron. **ME–6,** 167 (1959).

Determination of starch in paper. A comparison of the TAPPI, enzymatic, and spectrophotometric methods, J. L. Harvey, B. W. Forshee, and D. G. Fletcher, Tappi **42,** No. 11, 878 (1959).

Young's modulus of various refractory materials as a function of temperature, J. B. Wachtman, Jr., and D. G. Lam, Jr., J. Am. Ceram. Soc. **42,** No. 5, 254 (1959).

Branched-chain higher sugars. III. A 4-C-(hydroxymethyl)-pentose, R. Schaffer, J. Am. Chem. Soc. **81,** 5452 (1959).

Vibrational spectrum of cyanate ion in various alkali halide lattices, A. Maki and J. C. Decius, J. Chem. Phys. **31,** No. 3, 772 (1959).

Glass research at the National Bureau of Standards, C. H. Hahner, The Glass Industry, p. 588 (1959).

Applications of the molecular refractivity in radio metrology, B. R. Bean and R. M. Gallet, J. Geophys. Research **64,** No. 10, 1439 (1959).

Use of crystal to display high energy X-ray images, J. S. Pruitt, J. Soc. Non-Destructive Testing **XVIII,** No. 6, 359 (1959).

The measuring process, J. Mandel, Technometrics **1,** No. 3, 251 (1959).

Thermodynamic properties of helium at low temperatures and high pressures, D. B. Mann and R. B. Stewart, J. Heat Transfer **81,** 323 (1959).

Fire research at the National Bureau of Standards, A. F. Robertson, Fire Research Abstr. Rev. **I,** No. 4, 159 (1959).

Refractive indices and transmittances of several optical glasses in the infrared, G. W. Cleek, J. J. Villa, and C. H. Hahner, J. Opt. Soc. Am. **49,** No. 11, 1090 (1959).

Absolute photometry of the aurora—I. The ionized molecular nitrogen emission and the oxygen green line in the dark atmosphere, M. H. Rees, J. Atmospheric and Terrest. Phys. **14,** 325 (1959).

Absolute photometry of the aurora—II. Molecular nitrogen emission in the sunlit atmosphere, M. H. Rees, J. Atmospheric and Terrest. Phys. **14,** 338 (1959).

Preliminary assessment of the IGY, A. H. Shapley, Proc. Nat. Electron. Conf., p. 1 (1958).

Factorial experiments in life testing, M. Zelen, Technometrics **1,** No. 3, 269 (1959).

Direction findings on whistlers, J. M. Watts, J. Geophys. Research **64,** No. 11, 2029 (1959).

The construction of Hadamard matrices, E. C. Dade and K. Goldberg, Michigan Math. J. **6,** 247 (1959).

Geomagnetic and ionospheric phenomena associated with nuclear explosions, S. Matsushita, Nature **184,** BA33 (1959).

Stress corrosion cracking in low carbon steel, H. L. Logan, Proc. Phys. Met. Stress Corrosion Fracture Conf. Pittsburgh, Pa., Apr. 2–3, 1959, reprinted from Met. Soc. Conf., p. 295 (Interscience Publishers, New York, N.Y., 1959).

Higher oxides of silver, H. E. Swanson, J. Electrochem. Soc. **106,** No. 12, 1073 (1959).

Oscillatory behavior of magnetic susceptibility and electronic conductivity, A. H. Kahn and H. P. R. Frederikse, vol. IX, Adv. Solid State Phys., p. 257 (Academic Press Inc., New York, N.Y., 1959).

Geomagnetic activity following large solar flares, C. S. Warwick and R. T. Hansen, J. Atmospheric and Terrest. Phys. **14,** 287 (1959).

Bending and stretching of corrugated diaphragms, R. F. Dressler, ASME Trans. **81D,** No. 4, 651 (1959).

Simplification of systems of units currently evolving in the electrical field, F. B. Silsbee, Systems of Units, Am. Assoc. Advance. Sci., p. 7 (1959).

Atoms and free radicals by γ-irradiation at 4.2° K, L. A. Wall, D. W. Brown, and R. E. Florin, J. Phys. Chem. **63,** 1762 (1959).

Rotational transfer in the fluorescence spectrum of $OH(^2\Sigma^+)$, T. Carrington, J. Chem. Phys. **31,** No. 5, 1418 (1959).

Determination of viscosity, J. F. Swindells, ch. XII of Physical methods of organic chemistry, Part 1, p. 689 (Interscience Publishers Inc., New York, N.Y., 1959).

Triaxial tension at the head of a rapidly running crack in a plate, J. M. Frankland, Trans. ASME, Ser. E **26,** No. 4, 570 (1959).

Poly(tetrafluorethylene)—a radiation-resistant polymer, L. A. Wall and R. E. Florin. Letter to Editor, J. Appl. Polymer Sci. II, No. 5, 251 (1959).

On the approximate daytime constancy of the absorption of radio waves in the lower ionosphere, S. Chapman and K. Davies, J. Atmospheric and Terrest. Phys. **13,** No. 1–2, 86 (1958).

An X-ray crystallographic study of Schroeckingerite and its dehydration product, D. K. Smith, Am. Mineralogist **44,** 1020 (1959).

Modular forms whose coefficients possess multiplicative properties, M. Newman, Ann. Math. **70,** 478 (1959).

A note on the scattering of electrons from atomic hydrogen, A. Temkin, Phys. Rev. **116,** No. 2, 358 (1959).

The analysis of lactic squares with a certain type of row-column interaction, J. Mandel, Technometrics **1,** No. 4, 379 (1959).

Design of free-air ionization chambers for the soft X-ray region (20–100 kv), V. H. Ritz, Radiology **73,** No. 6, 911 (1959).

Thermal conductivity of indium antimonide at low temperatures, E. V. Mielczarek and H. P. R. Frederikse, Phys. Rev. **115,** 888 (1959).

Model for vibrational relaxation of diatomic gases behind shock waves, R. Herman and R. J. Rubin, Phys. of Fluids **2,** No. 5, 547 (1959).

A note on algebras, A. J. Goldman, Am. Math. Mo. **66,** 795 (1959).

Transistor P–A amplifier, G. F. Montgomery and F. R. Bretemps, Electronic Ind. 19, No. 1, 196 (1960).

Linear differential equations of the second order with a large parameter, F. W. J. Olver, J. Soc. Indus. Appl. Math. **7,** 306 (1959).

Sorting devices in postal tests. 1: Proposal for an electromagnetic sorter, S. Henig, Automatic Control **11,** No. 6, 6 (1959).

The dissociation constant of $CaOH^+$ from 0° to 40° C, R. G. Bates, V. E. Bower, R. G. Canham, and J. E. Prue, Trans. Faraday Soc. **55,** No. 444, 2062 (1959).

Evaluation of chemical analyses on two rocks, W. J. Youden, Technometrics **1,** No. 4, 409 (1959).

Nuclear electronics, L. Costrell, Nucleonics Mag. **18,** No. 1, 124 (1960).

Microscopical studies of failure in polymers, S. B. Newman, ASTM Spec. Tech. Pub. **257,** 132 (1959).

Flare-associated bursts at 18 Mc/s, C. Warwick and J. W. Warwick, Paris Symp. on Radio Astron. (LAU Symp. No. 9 and URSI Symp. No. 1), p. 203 (1959).

Quantum-mechanical calculation of the probability of an exchange reaction for constrained linear encounters, J. Mazur and R. J. Rubin, J. Chem. Phys. **31,** No. 5, 1395 (1959).

Monte Carlo calculations of gamma ray backscattering, M. J. Berger and D. J. Raso, Radiation Research **12,** No. 1, 20 (1960).

Identification of textile coatings by infrared spectroscopy, F. H. Forziati, R. T. Hite, and M. K. Wharton, Am. Dyestuff Reptr. **49,** No. 4, 29 (1960).

New wavelengths for some helium (He I) lines, W. C. Martin, J. Opt. Soc. Am. **50,** No. 2, 174 (1960).

Geomagnetic effects of high-altitude nuclear explosions, A. G. McNish, J. Geophys. Research **64,** No. 12, 2253 (1959).

Fading rate recorder for propagation research, J. W. Koch, W. B. Harding, and R. J. Jansen, Electronics **32,** No. 51, 78 (1959).

Matrix algebra for calculating multicomponent mixtures, F. Ordway, Portland Cement Assoc. J. Research and Devel. Labs. **2,** No. 1, 28 (1960).

Determination of piezoelectric properties as a function of pressure and temperature, J. E. McKinney and C. S. Bowyer, J. Acoust. Soc. Am. **32,** No. 1, 56 (1960).

Line shape and f value in the $OH^2\Sigma^+ - {}^2\pi$ transition, T. Carrington, J. Chem. Phys. **31,** No. 5, 1243 (1959).

Determination of the recording performance of a tape from its magnetic properties, E. D. Daniel and I. Levine, J. Acoust. Soc. Am. **32,** No. 2, 258 (1960).

Leonard Euler's integral: A historical profile of the gamma function, P. J. Davis, Am. Math. Mo. **66,** 849 (1959).

Gages for measuring the thickness of chromium on the internal surface of small-bore tubes, V. A. Lamb and P. A. Krasley, Plating **47,** No. 2, 176 (1960).

Departures from the Saha equation under varying conditions of Lyman continuous opacity, S. R. Pottasch and R. N. Thomas, Astrophys. J. **130,** No. 3, 941 (1959).

Experimental and theoretical investigation of the magnetic properties of iron oxide recording tape, E. D. Daniel and I. Levine, J. Acoust. Soc. Am. **32,** No. 1, 1 (1960).

Vibration-rotation bands of ammonia. IV. The stretching fundamentals and associated bands near 3μ, W. S. Benedict E. K. Plyler, and E. D. Tidwell, J. Chem. Phys. **32,** No. 1, 32 (1960).

Electron impact studies of aromatic hydrocarbons, I. Benzene, napthalene, anthracene, and phenanthrene, M. E. Wacks and V. H. Dibeler, J. Chem. Phys. **31,** No. 6, 1557 (1959).

Value of the Rydberg constant, W. C. Martin, Phys. Rev. **116,** No. 3, 654 (1959).

Characteristics of deposits, W. H. Metzger, Jr., Symp. Electroless Nickel Plating, ASTM Spec. Tech. Pub. **265,** p. 13 (1959).

Turbulent motion, G. B. Schubauer and C. M. Tchen, sec. B, vol. V, Princeton Ser., High speed aerodynamics and jet propulsion, Turbulent flows and heat transfer, p. 75 (Princeton Univ. Press, Princeton, N. J., 1959).

Confidence intervals for the expectation of a Poisson variable, E. L. Crow and R. S. Gardner, Biometrika **46,** 441 (1959).

Electron characteristic energy losses in some intermetallic compounds, B. Gauthe, Phys. Rev. **114,** No. 5, 1265 (1959).

Adsorption, diffusion, and evaporation of carbon monoxide or tungsten, R. Klein, J. Chem. Phys. **31,** No. 5, 1306 (1959).

Efficient harmonic generation, G. F. Montgomery, Proc. IRE **48,** No. 2, 251 (1960).

Effect of fluorides on infrared transmittance of certain silicate glasses, G. W. Cleek and T. S. Scuderi, J. Am. Ceram. Soc. **42,** No. 12, 599 (1959).

Studies of borate minerals. VI: Veatchite, J. R. Clark, M. E. Morse, A. Perloff, and G. Burley, Am. Mineralogist **44,** No. 11–12, 1141 (1959).

Conditions at the ionization and shock fronts in collisions of gas clouds—Bright rims in diffuse nebulae, pt. VI, S. R. Pottasch, Rev. Mod. Phys. **30,** No. 3, 1053 (1958).

The adhesion of electrodeposited nickel to chromium at elevated temperatures, W. E. Reid, Jr., and F. Ogburn, J. Electrochem. Soc. **107,** No. 2, 91 (1960).

Chloride content of the diffusion layer at a silver anode, P. A. Krasley, J. Electrochem. Soc. **107,** No. 2, 139 (1960).

Spectral emittance of ceramic-coated and uncoated specimens of Inconel and stainless steel, J. C. Richmond and J. E. Stewart, J. Am. Ceram. Soc. **42,** No. 12, 633 (1959).

Mismatch errors in cascade-connected variable attenuators, G. E. Schafer and A. Y. Rumfelt, IRE Trans. Microw. Theory Tech. MTT–**7,** No. 4, 447 (1959).

Chemical structure and stability relationships in polymers, B. G. Achhammer, M. Tryon, and G. M. Kline, Kunststoffe combined with German Plastics Digest **49,** No. 11, 600 (1959); Modern Plastics **37,** No. 4, 131 (1959).

A study of 17–7 *p*H stainless steel, N. L. Carwile and S. J. Rosenberg, WADC Tech. Rep. 58–653 (1959), available from the Office of Technical Services, Department of Commerce, Washington 25, D.C.

**Publications for which a price is indicated (except for NBS Technical Notes) are available only from the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. (foreign postage, one-fourth additional). Technical Notes are available only from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D.C. (Order by PB number.) Reprints from outside journals and the NBS Journal of Research may often be obtained directly from the authors.*

Notice

The papers listed below will appear in early issues of the Journal of Research, Section D, Radio Propagation:

ELF electric fields from thunderstorms. A. D. Watt.

Field strength measurements in fresh water. Gurdip S. Saran and Gedalia Held.

Electrical resistivity studies on the Athabasca Glacier, Alberta, Canada. George V. Keller and F. C. Frischknecht.

The amplitude distribution for radio signals reflected by meteor trails. A. D. Wheelon.

Computation and measurement of the fading rate of moon-reflected UHF signals. S. J. Fricker, R. P. Ingalls, W. C. Mason, M. L. Stone, and D. W. Swift.

On the theory of wave propagation through a concentrically stratified troposphere with a smooth profile. H. Bremmer.

Polarization and depression-angle dependence of radar terrain return. I. Katz and L. M. Spetner.

Methods of predicting the atmospheric bending of radio rays. B. R. Bean, G. D. Thayer, and B. A. Cahoon.

Loss in channel capacity resulting from starting delay in meteor-burst communication. George R. Sugar.

Elementary considerations of the effects of multipath propagation in meteor-burst communication. George R. Sugar, Robert J. Carpenter, and Gerard R. Ochs.

Use of logarithmic frequency spacing in ionogram analysis. G. A. M. King.

The guiding of whistlers in a homogeneous medium. R. L. Smith.

The propagation of microwaves through a magneto-plasma, and a possible method for determining the electron velocity distributions. A. L. Cullen.

On electromagnetic radiation in magneto-ionic media. Herwig Kogelnik.

Radiation and admittance of an insulated slotted-sphere antenna surrounded by a strongly-ionized plasma sheath. John W. Marini.

A contribution to the theory of corrugated guides. Gerhard Piefke.

High-gain, very low side-lobe antenna with capability for beam slewing. A. C. Wilson.

Shielding of transient electromagnetic signals by a thin conducting sheet. Norman Zitron.

Cylindrical antenna theory. R. H. Duncan and F. A. Hinchey.

UNITED STATES
GOVERNMENT PRINTING OFFICE
DIVISION OF PUBLIC DOCUMENTS
WASHINGTON 25, D.C.

OFFICIAL BUSINESS

PENALTY FOR PRIVATE USE TO AVOID
PAYMENT OF POSTAGE, $300
(GPO)

JOURNAL OF RESEARCH of the National Bureau of Standards Vol. 64D, No. 4 July–August 1960
D. Radio Propagation

Contents

	Page
Relation of turbulence theory to ionospheric forward scatter propagation experiments. Albert D. Wheelon.	301
Propagation at oblique incidence over cylindrical obstacles, M. P. Bachynski.	311
Diffraction by smooth conical obstacles. H. E. J. Neugebauer and M. P. Bachynski.	317
Characteristics of 488 megacycles per second radio signals reflected from the moon. B. C. Blevis and J. H. Chapman.	331
The use of polarization fading of satellite signals to study the electron content and irregularities in the ionosphere. C. Gordon Little and Robert S. Lawrence.	335
Note on a test of the equivalence theorem for sporadic *E* propagation. J. W. Wright and T. N. Gautier.	347
Daytime attenuation rates in the very low frequency band using atmospherics. W. L. Taylor.	349
Measured electrical properties of snow and glacial ice. A. D. Watt and E. L. Maxwell.	357
Half-wave cylindrical antenna in a dissipative medium: current and impedance. Ronold King and Charles W. Harrison.	365
Preface to ELF papers.	381
Some ELF phenomena. E. T. Pierce.	383
Mode theory and the propagation of ELF radio waves. James R. Wait.	387
Studies of natural electric and magnetic fields. G. D. Garland and T. F. Webster.	405
Natural electromagnetic energy below the ELF range. Wallace H. Campbell.	409
Possible application of the system loss concept at ELF. Kenneth A. Norton.	413
Measurements of the spectrum of radio noise from 50 to 100 cycles per second. M. Balser and C. A. Wagner.	415
Listing of published VLF symposium papers.	419
Publications of the staff of the National Bureau of Standards.	420

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington 25, D.C. Price 70 cents (single copy).
Subscription price: $4.00 a year; $0.75 additional for foreign mailing.

UNITED STATES GOVERNMENT PRINTING OFFICE, WASHINGTON : 1960

opy).